

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: _____ Examiner #: _____ Date: _____
Art Unit: _____ Phone Number 30 _____ Serial Number: _____
Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

For Sequence Searches Only Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

Jan Delaval
Reference Librarian
Biotechnology & Chemical Library
CM1 1E07 - 703-308-4498
jan.delaval@uspto.gov

STAFF USE ONLY

	Type of Search	Vendors and cost where applicable
Searcher: <u>Jan</u>	NA Sequence (#) _____	STN _____
Searcher Phone #: <u>448</u>	AA Sequence (#) <input checked="" type="checkbox"/> _____	Dialog _____
Searcher Location: _____	Structure (#) _____	Questel/Orbit _____
Date Searcher Picked Up: <u>4/23/03</u>	Bibliographic _____	Dr.Link _____
Date Completed: <u>4/23/03</u>	Litigation _____	Lexis/Nexis _____
Searcher Prep & Review Time: _____	Fulltext _____	Sequence Systems <input checked="" type="checkbox"/> _____
Clerical Prep Time: <u>10</u>	Patent Family _____	WWW/Internet _____
Online Time: <u>10</u>	Other _____	Other (specify) _____

Result No.	Score	Query Match	Length	DB	ID	Description
1	3033.5	72.4	803	2	B56272	DNA-directed DNA I
2	2434	58.1	784	2	E72515	Probable DNA-directed
3	2246	53.6	781	2	JC7382	DNA-directed DNA I
4	1581	37.8	781	2	A69312	DNA polymerase B1
5	1287	30.7	764	2	S75477	Probable DNA-directed
6	1236.5	29.5	775	2	S35543	DNA-directed DNA I
7	1207.5	28.8	771	2	C75023	DNA polymerase I
8	1197.5	28.6	775	2	S67930	DNA-directed DNA I
9	965.5	23.1	1312	2	S68593	DNA-directed DNA I
10	955.5	22.8	1235	2	C71210	Probable DNA-directed
11	915	22.8	586	2	C69028	DNA-dependent DNA
12	771.5	18.4	1086	2	T40282	DNA polymerase de
13	768.5	18.4	1670	2	S71551	DNA-directed DNA I
14	763.5	18.2	1086	2	T43266	DNA-directed DNA I
15	760.5	18.2	1107	2	A41618	DNA-directed DNA I
16	757.5	18.1	1106	1	A39299	DNA-directed DNA I
17	751	17.9	1038	1	UC5757	DNA-directed DNA I
18	749	17.9	1038	2	T18222	DNA-directed DNA I
19	743.5	17.8	1105	1	S40243	DNA polymerase de
20	741.5	17.7	1088	2	T05721	DNA-directed DNA I
21	731.5	17.5	1084	1	S19661	DNA-directed DNA I
22	718.5	17.2	901	2	E84210	DNA polymerase B1
23	713	17.0	1702	2	S82429	DNA-directed DNA I
24	705	16.8	1094	2	S22573	DNA-directed DNA I
25	691.5	16.5	1097	1	RNBV13	DNA-directed DNA I
26	671	16.0	1634	0	BE6410	DNA-directed DNA I
27	649.5	15.5	879	2	A56277	DNA-directed DNA I
28	647.5	15.5	872	2	JC7380	DNA-directed DNA I
29	628.5	15.0	875	2	UC5166	DNA-directed DNA I

30	625.5	14.9	959	2	F72766	probable DNA-directed
31	598.5	14.3	882	2	T23019	DNA-directed DNA p
32	598	14.3	1081	2	T20698	hypothetical prote
33	594.5	14.2	882	2	F90201	DNA polymerase I (
34	579.5	13.8	1462	1	D4HUAC	DNA-directed DNA p
35	578	13.8	1465	2	S45628	DNA-directed DNA p
36	569	13.6	787	2	E83227	DNA-directed DNA p
37	564.5	13.5	844	2	T31321	DNA polymerase II
38	527.5	12.6	787	2	G83410	DNA-directed DNA p
39	526	12.6	3122	2	T17202	DNA polymerase II
40	524.5	12.5	1015	1	D1BE2L	DNA-directed DNA p
41	522.5	12.5	1339	1	S20052	DNA-directed DNA p
42	517	12.3	1505	2	S28079	DNA-directed DNA p
43	515	12.3	1009	1	D3EBM2	DNA-directed DNA p
44	511	12.2	1026	2	T03108	DNA-directed DNA p
45	508.5	12.1	1513	2	T25158	probable DNA-direct

ALIGNMENTS

5A No

RESULT 1
B56277
DNA-directed DNA polymerase (EC 2.7.7.7) II - Pyrodicticum occultum
C:Species: Pyrodicticum occultum
C:Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #ext_change 20-Jun-2000
C:Accession: B56277
R:Uemori, T.; Ishino, Y.; Doi, H.; Kato, I.
J. Bacteriol. 177, 2164-2177, 1995
A:Title: The hyperthermophilic archaeon Pyrodicticum occultum has two alpha-like DNA polymerases
A:Reference number: A56277; PMID:95238290; PMID:7721707

A:Accession: B56277
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-803 <DEM>
A:Cross-references: GB:D38574; NID:g807829; PIDN:BA407360.1; PID:g807830
C:Superfamily: herpesvirus DNA-directed DNA polymerase
/keywords: nucleotidyltransferase

Query Match	72.4%;	Score 3033.5;	DB 2;	Length 803;
Best Local Similarity	70.9%;	Pred. No. 1.4e-176;		
Matches 571;	Conservative 106;	Mismatches 121;	Indels 7;	Gaps 5

[illegible]

Db 419 YGAVVLLKPLKGVHENVVLDLDFSSMPSIMIKYVGPDTIVDDPSECEPKYGGCYVAPEVG 478
 Qy 477 HFRRCPPGPFKTVLERLELRLKRVABEMKKYPPDSPEYRLDLDEROKALKVLANASYGM 536
 Db 479 HFRRCPPGPFKTVLERLELRLKRVABEMKKYPPDSPEYRLDLDEROKALKVLANASYGM 538
 Qy 537 GMSGARVYCRECAKAVTAMGRHLIRTAIINARKGLKTVYGDYDLSLFTVYDPEKVENFLK 596
 Db 539 GMSHARVYCKRCACAVTAMGRHLIRTAIINARKGLKTVYGDYDLSLFTVYDPEKVENFLK 598
 Qy 597 IKKELEBEITLERNVYKRLPPTAKKRYAGLLEDERIDIVGFVAVNGDMCELAKEVQTKV 656
 Db 599 FVEKELGEIKIDIKYKVFTEAKKRYVGLLEDERIDIVGFVAVNGDMCELAKEVQTKV 658
 Qy 657 VEIVLKTSEVNAKAVEYRKIVKELEKVPLEKVIYKTLSEYTTTEAPHVVAARQM 716
 Db 659 AEIVLNTGVNDKAIISYIREVYIKQIREGKVPITKLIWTKLSKRIEYEHADAPVMAARM 718
 Qy 717 LSAGYVSPGDKIGYVIVKGGRIISQANPFYVNDPSQIDVTYVDHQTIPALRIILGY 776
 Db 719 KEAGYEVSPGDKIGYVIVKGGSGSVSRAYPFYFMV-DPSTIDVNYIIDHQTIPALRIILSY 777
 Qy 777 FGITEKLLKASATGQKTLPDFLAKK 801
 Db 778 FGITEKOLKAAATVQSLPDFPASK 802

RESULT 2

probable DNA-directed DNA polymerase APE2098 - Aeropyrum pernix (strain K1)
 C:Species: Aeropyrum pernix
 C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
 C:Accession: E72515
 R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Halkawa, Y.; Jin-no, K.; Takah
 ame, H.; Takamaya, M.; Maeda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
 DNA Ref. 6, 83-101, 1999
 A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
 A:Accession: E72515
 A:Reference number: A72450; M0ID:99310339; PMID:10382966
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-784 <KAM>
 A:Cross-references: DDBJ:AB000663; NID:95105654; PIDN:BAAB1109.1; PID:95105797
 A:Experimental source: strain K1
 C:Genetics:
 A:Gene: APE2098
 C:Superfamily: herpesvirus DNA-directed DNA polymerase

Query Match 58.1%; Score 2434; DB 2; Length 784;
 Best Local Similarity 59.1%; Pred. No. 3 7e-140;
 Matches 469; Conservative 113; Mismatches 189; Indels 22; Gaps 5;

Qy 17 GCEPQVITMGIAENGERVYLIDSRFPFYALLAPGADPKQVACORIALSRPSPIIGYE 76
 Db 3 GSTPITIMGAGDSKRVVFTGEFRPFYVLPDSYGLDQLAMIRLRSPSSPLISYE 62
 Qy 77 DDKKRYGRRPRVLRIRTVLPEAVREYRELKNDVGEVYLEADIRFARKYILIDHLPF 136
 Db 63 RRRRPFIGREYALKVTVLPASVREYRAVARLOGVADLEADIPALRFIIDIRLYPM 122
 Qy 137 TWYVEAEPLLENKMGFRVDKVVY-----KSREPLYGEALATPKLPDLRIIADI 187
 Db 123 RMYVAEVEVAVPHGYSDVDAVYLSGDIRREDETRIOEDPLKG-----LRVAFDI 172
 Qy 188 EYVSKQSGPRERDPVIVIAVKTDGDEVLFIAEGDKDRKPIREFVEYVYKVDPIITIGY 247
 Db 173 EYVSKMRTPPDKOPVIMIGLQOAGGEIRILEAEDRSKVIAGFVERKSIDPVIYGY 232
 Qy 248 NNNHDPMPYLLRARIIGLDVTRVGAPEPTSHGVSPGLAVNDLYDAEEMPEIK 307
 Db 233 NQNRDMPYLVERRARVLGVTLAVGR-SVEPGLXGHYSVSGRLVNDLDPREELHEVK 291
 Qy 308 IKSLEVAEYVIGVWKKSERVIINWMEIPYNDPKKRPLLIQVARDVDVATYGLAEKILP 367

Db 292 VKTLEEVADYGVNVIGERVYLTLEWQIGEWDPDKREILIKRYLRDDVRSTMGIAKFP 351
 Qy 368 FALQSLYVNGPLPDQVAMSVGRLEWYLLIRAPFMKELVNRYVERDEEYRAIYLEL 427
 Db 352 FGAEISQVSGPLPDQVAMSVGRLEWYLLIRAPFMKELVNRYVERDEEYRAIYLEL 411
 Qy 428 RGHENIANTDSSSTPNIMIKTVNGPDTLVNPGKCECCGMAPEYKIRPFCPPGF 487
 Db 412 PGHEIDIALDASMYPIWIKTVNGPDTLVNPGKEVEEYVAPPEYKIRPFCPPGF 471
 Qy 488 KTVLEELLERKRVABEMKKYPPDSPEYRLDLDEROKALKVLANASYGMGSGARVYCE 547
 Db 472 KTLERFLSMRQIRSEKKGPPDSPEYRLDLDEROKALKVLANASYGMGPHARVYCRE 531
 Qy 548 CAAVATAMGRHLIRTAIINARKGLKTVYGDYDLSLFTVYDPEKVENFLIKKELEGEIK 607
 Db 532 CAEAVTAMGRSLIRTAIRKAGELEIYIGDYDLSLFTVNDPEKVERLIRFVEBELGFDIK 591
 Qy 608 LEKVYRLPFTFAKKRYAGLLEDERIDIVGFVAVNGDMCELAKEVQTKVVEIVLKTSEV 667
 Db 592 VDKVRYVFTFAKKRYVGLVDQKIDVGFVAVNGDMSELAKEVQTKVVEIVLKTSGVD 651
 Qy 668 KAVEYRKIVKELEKGVPIEKVIYKTLSEYTTTEAPHVVAARQMISAGYVSPGD 727
 Db 652 EADVYRNRIEKLARGOVDMRKVIYKTLIRPSRYEARQHVYVALLMERAGIKVEPBA 711
 Qy 728 KIGYVIVKGGRIISQANPFYVNDPSQIDVTYVDHQTIPALRIILGYITEKLLKAS 787
 Db 712 KIGYVIVKGGSGSVSRAYPFYFMV-SKEVDEYVYDKOVVPALRIILGYITEKLLKG 770
 Qy 788 ATGQKTLPDFLAKK 800
 Db 771 GR-QSTLIDPMKR 782

RESULT 3

DNA-directed DNA polymerase (EC 2.7.7.7) B3 - Sulfurleptaera ohkawensis
 N:Alternate names: DNA polymerase alpha, beta, gamma, DNA polymerase I, II, III
 C:Species: Sulfurleptaera ohkawensis
 C:Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 31-Dec-2000
 C:Accession: J07382
 R:Iwai, T.; Kurosawa, N.; Itoh, Y.H.; Kimura, N.; Horuchi, T.
 DNA Ref. 7, 243-251, 2000
 A:Title: Sequence analysis of three family B DNA polymerases from the thermophilic
 A:Reference number: J07380
 A:Accession: J07382
 A:Molecule type: DNA
 A:Residues: 1-781 <IMA>
 A:Cross-references: DDBJ:AB032376
 C:Comment: This enzyme has both 3'-5' exonuclease and polymerase activities, and plays
 C:Genetics:
 A:Gene: B3
 C:Keywords: DNA replication; exonuclease; metal binding; nucleotidyltransferase

Query Match 53.6%; Score 2246; DB 2; Length 781;
 Best Local Similarity 54.3%; Pred. No. 9.9e-129;
 Matches 432; Conservative 141; Mismatches 203; Indels 20; Gaps 10;

Qy 6 FTVLDSYEVNGKEPQVITMGIAENGERVYLIDSRFPFYALLAPGADPKQVACORIAL 65
 Db 5 FTLDPSYDVVENKRPVITWIDKGNRVLLLEKFRFYFALVDSDVINDEIRKELVL 64
 Qy 66 SRKPSITIGVDDDKRKYGRRPRVLRIRTVLPEAVREYRELKNDVGEVYLEADIRF 125
 Db 65 SKRPSITISIDVEKKYFGSPVAVLKIEIVYIAYRVYRDVAIKGVKSLVLEADIRFM 124
 Qy 126 RYILIDHLPFTPTVTRVVAEPLLENKMGFRVDKVVYKSPPELYGEALATPKLPDLRII 185
 Db 125 RISIDINLPFTPTVTRVVAEVEELKEN-NFRVYKVELEK-KINLVYED-----KIPELVLA 177
 Qy 186 DIEVSKQSGPRERDPVIVIAVKTDGDEVLFIAEGDKDRKPIREFVEYVYKVDPIIT 245


```

Db 178 DIEVNMKXGSPNRPDDPIIIGVWTKBGGK-QFLADKIDDLRAHEFINFQYTDPIIV 236
Qy 246 GYNNHFDMPYILRRPARILIGIKLDVTRVGAEPPTSVHGHSVPGRLNVDLYDAEMPE 305
Db 237 GYNINFPMPYILREBANIIGIRLIDVGRNNGEPEGSQVGHYSITGRNLNVLYGFAQIOE 296
Qy 306 IKIKSLEEVAELGYMKKSERVIINMWEPDYMDPKRPRILLQVARDPDAATGLAEKI 365
Db 297 VKVKLTENIADTSLVPEKRTIVEMYDIPKWDDEKRDILKTNLDDAKSAYILGEVF 356
Qy 366 LPFAIQLSVTGLPIDOVGAMSVEGRLEWYILRAAFKKELVPRVERPEETYGAIYLE 425
Db 357 IPFGIELTRISGLPIDOLSMASVGHVRVEMLMREBAVKYNNELIPNKEERESEYEGALVIS 416
Qy 426 PLRGYHEINAVLDPSSMPTNIMIKYNGPDTLYRGEKCGEC-CGWEAPDYKHRRPRPP 484
Db 417 PLRGHEEYVYDPSMPTSMIKINIGDILVK----GGEENCWSP-VGHRKREPP 470
Qy 485 GFETVLERLLELRKRVRAEMKKYPPDSPERYLLDEROKALKVLNNAASYGMWGSARWY 544
Db 471 GLYKVVLEKTLQERKEVKYKLMK-TIDEYDKRVLDAORALKVMANA FYGMWGLGARWY 529
Qy 545 CRECAKATWAMGRHLIRTAINIRAKLGLKVIYGDPIISLPMYDPEKVNPKIKIKEELGF 604
Db 530 SKEGEBAATAMORQIISDSAKTAKEGFTVIYGDPIISFVKGSG-INSLTLESKFG 588
Qy 605 EIKLEKVKRLLEFTEAKKRYAGLLEDGRIDIVGEFAVBDWCCELAKEVOTKVEIVLTKTS 664
Db 589 EIKIKIKYKRVFPTENKKRYAGLTEDGKIDIVGEFAVBDWCIDLAKQVOTNIVELILKSG 648
Qy 665 EYNKAVEVYRKIVKLEBEGKVPITELVIMKTLSKRLEEYTBAPHVVAAKMLSGRYVS 724
Db 649 KWEDAIKTVKSYIFPLRRYNFRIDELIIMTKYDKLDEVDYTAHPVVAAKKAARAGLYVS 708
Qy 725 PDKKIGYIVVKSGGRISQRAWPFVNWQPSQIDVYVYUDHOIIPALAILIGFGITKEKL 784
Db 709 KGVKIGYIVVKSGKISDRAEPYFLVKEKNKIDIVEYIYIDKOIIPALAILIGFGVKESSL 768
Qy 765 KASATGQKTLPPDLAK 800
Db 769 K--TGSVDILSFPKK 781

```

RESULT 4

DNA polymerase B1 (polB) homolog - Archaeoglobus fulgidus
Species: Archaeoglobus fulgidus
Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
Accession: A69312
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
J.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kittneiss, E.F.;
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Uitterlisch, T.; Cotton, M.D.; Spriggs, T.; Attiach, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Moese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A:Reference number: A69250; MUID:98049343; PMID:9389475
A:Accession: A69312
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-781 <KLE>
A:Cross-References: GB:AE001070; GB:AE000782; NID:G2689393; PIDN:AA90741.1; PID:G265013
C:Superfamily: herpesvirus DNA-directed DNA polymerase

	Query Match	37.8%	Score 1581;	DB 2	Length 781;
	Beat Local Similarity	42.5%	Pred. No. 2,6e+88;		
	Matches 350; Conservative	133;	Mismatches 279;	Indels 68;	Gaps 18;
QY	1 MREVFVTVDDSSYEVGKREPQIVTINCIAENGERVVLIDSPRRPYRATALLAPGADPQYAQ	60			
	:::::	:	:	:	:
Db	1 MRERVGMLLDAIDDETGGKAAYVRLMKCKDQG--IIPAAVDNFPDPFVFI--GVSDILIKN	56			

```

0Y 61 RIALSRKSPKSP1IGVEDDKKRYFGRRPRVLRIRVYLPBAAVEYELVKNNDGVEDVLEAD 120
Db 57 AATSRREVRVILKSKSEPKQOLKTLGRVEGVIVYAHNPHOVHFKLDLYLSQFG--DVREAD 113
0Y 121 IFAARVYILDDLPFPFTYRVEABPLENKG-----PRVDKVLVKSAPPELYGEALAPT 175
Db 114 IFAATVYILDDKDLACMDGIAIEGE-----KQGVIRSKYIEKV-----ERIPRM 157
0Y 176 KL.PDLIRILAFDIEVYSKQSPRPERDPVYIVAVXTDDDEVLFIAEGDKDRKPIREFVEY 235
Db 158 EEPFLKMLVPCDEMISFGMEPEKEDPIVIVSXTNDDDEILING--DEKKIISDFVL 214
0Y 236 VKRYPDIIIVGNNNNHPWPIILRRARILGIKLDVTRRVCAGLEPTTSH--GHVSVPGELN 293
Db 215 IKSYPDIIIVGNDAPFMPYILRRKAERMNPLD---VGDGGSNVVFRGGRPIITGRLN 270
0Y 294 VDLVYAEEMPEIKI KSJEEVAEYLGVMKKSERYINMMELPDVWDMDPKKPLLOYARD 353
Db 271 VDLVDIARISDIIKILGLENVAEFLGT--KIEIADIEXKDIYRHSRGEKK--VLNARQ 327
0Y 354 DVRAIVYLAELKILPFAIQLSVYVGTPLDQVGAMSVGFLEWYILRAAFKMKELVPNRVER 413
Db 328 DAINIVYILAKELLPMHWYELSKMIRLPVDDVTRMRGKQVMDLLSEAKKIGELIAPNPREH 387
0Y 414 PEETRGVILVLEPLRGVHENTAVLDQSGMYPMNITKYNAGDPLVLRPEKKGEGCGMEAP 473
Db 388 -AESYEGFVLEPEPEGHLENVACUDFASMYPISIMAFNISDPTV---GGRD--DCYEAP 440
0Y 474 EVKHFRRACPGFPFTVLERLLELRKRVRAEMKKYPPSPERYLLDEROKMLKYLAAASY 533
Db 441 EVGHFERSPGOGFFRGIIRMLIEKRRELKVBKMLSPSSSYKLLDIIKOQLTKVLTNSFY 500
0Y 534 GYMGSGARWYCREBCAKAVTAMGSHLIRTAINTARKLGLKVIYSDPTSLFVT---YDPEK 590
Db 501 GYMGNILARWYCHPCAETATAMGRHFIITSAKIAESMGFKVLYGDTDSIFVTKAGMTKED 560
0Y 591 VENFIKIKKEELGPEIKLEKYYKRLFTEAKKRVAGLLDERIDIVGEAVNRGWMCELA 650
Db 561 VDLRLDKHEELPIQIEVEDEYISAFVE--KRRVAGLLEDGRLVVKGLEVRGWMCELA 619
0Y 651 EVQTKVEIVLKTSEVNNAVEYVRKIYKELEEGVPIEKLYIMKTLISKRLSEYTTAPHV 710
Db 620 KYQREIVLEILKKNPEKALSLAVKDVILRIIEGKVSLEEVYIYGGITKKPKSKESMAHV 679
0Y 711 VAAKMLSAGTRVSBGDKITGIYIVKGGRIISQRAWPFYFMVD-----PSQ 755
Db 680 KAALKAREMGIIYVSSKIGYIVIVGSGNIGDRAVPIDLLIEDFGENLRITKSGIEIKK 739
0Y 756 IDVTYVYDHOIIPALRILIGYFITEKKLKASAGQGLTLPPELA 799
Db 740 LKQYIIONQIIPSVLRIERFGTEASLKGS--QMSLDSFFS 781

```

RESULT 5

Probable DNA-directed DNA polymerase (EC 2.7.7.7) - *Sulfolobus solfataricus*
 N:Alternate names: protein c04041
 C:Species: *Sulfolobus solfataricus*
 C:Date: 11-Oct-1997 #sequence_revision 24-Oct-1997 #text_change 20-Jun-2000
 C:Accession: S755407
 R:Resener, C.W.; Klenk, H.P.; Singh, R.K.; Allard, G.; Chan, C.C.Y.; Liu, Q.Y.; Penny, S.
 Mol. Microbiol. 22, 175-191, 1996
 A:Title: Organizational characteristics and information content of an archaeal genome: 1
 A:Reference number: S73076; MUID:97055432; PMID:8899719
 A:Accession: S75407
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-764 <SEN>
 A:Cross-references: EMBL:Y08257; NID:g1707772; PID:g1707813
 A:Experimental source: strain P2
 A>Note: The nucleotide sequence was submitted to the EMBL Data Library, September 1996
 C:Superfamily: herpesvirus DNA-directed DNA polymerase
 C:Keywords: nucleotidyltransferase

Query Match 30.74; Score 1287; DB 2; Length 764;
 Best Local Similarity 34.84; Pred. No. 1.9e-70;
 Matches 280; Conservative 167; Mismatches 305; Indels 52; Gaps 16;

QY 6 FTVLDSYEVVNGKEPQVIIMGIAENGERVVLIDSPRPYFALLAGADPKQVAAQRRL 65
 D 5 FTLLDPSEYKNGTPELVYISVDENSSVYIDNNFRYFYIIE--GNNENIENI--- 59
 QY 66 SRKSPITIGVEDDKRYKFGRRPRVLRITVLEAVREYELVKVNDGVEDVLEADIRPM 125
 D 60 -KNGEALQITKVKKYLGINVDALLIGTSTFQIKKREKISLNNIKGIPADRIRYM 118
 QY 126 RYLIDHDLPEFTWYRVEAEPLNMGFRVDKYLKSRPPELYGEALAPTKLPDLRIALF 185
 D 119 RYSLDPDLRPFMTWFAEYNEVKFD-GFRTKKAYIIDKLISHYEG-----NMDELKTI 171
 QY 186 DLEVYSKQSPRPERDPIYIAVKTDDGDEVLFIAGKDDRKPIREVEYVKRPDPIY 245
 D 172 DQGIYSKQSLNPKRDPYVMSLSKSGPMQSLDEGIDDKIIRKRVYDIILATDPIIF 231
 QY 246 GYNNHFMPPYLBRARILGKLDVTRRGAEPTSVAGHVSFGRLANDLYAEEME 305
 D 232 VIDSDLPWKITITERBASLGKIDIGKIGSEVSGTGHYSISGRNLVDTGLVNER 291
 QY 306 IKIKLSEVAYELGVMMKSERVIINMWEIPDYDDPKKPLLOYADVRATYGLAEKI 365
 D 292 LGHVDLIDVSNLGI--SPSRYSFKWEISRYWDEKNNRIIREYSIENARSIYLLGNYL 349
 QY 366 LPPALQSLVYVGLPDDGVAMSVGFLEWYLIRAAFKMKELVNRVERPE--ETRGALV 423
 D 350 LSTYSELVYIVGLPLDKLSVASWGNRIETSLIRATKSGELIPTRMNPFRPSKIKNI 409
 QY 424 LEPARGHENTIAVLDFSSMYPNIMIKYNVGPDTLVREKSGCEGAEAVGHRPRCP 483
 D 410 IOPKGIYDYVDVLDISSVSLVRKENIAPDLVK--EQCDD--CYSSPISMYKREPR 465
 QY 484 PGFFTVLERLLELRKVRAMKYPDPDSPEYRLDEROKALKYLANASGYMGWGA 543
 D 466 SGLYKTFDELSNVR-----DSMKIKYIE-----LISFNDYHWANAM 506
 QY 544 YCECAKAVTAMGRLIRTAINTARKLGKVIYGDPSLFT--YDPRKNFPIKIKEL 602
 D 507 YSREIASAFDESNBIIRFIIDIKSSGDIYLANDLIPTGSRKKNBELTKINSIY 566
 QY 603 GPEIKLEKYKLEPTEKKRYAGLEDGRIDIVGFPAVRD---WCEIAKVOQRYVE 658
 D 567 NIDVAKTIFYKSLVLD--NNRIAGLSEGDID---ARKSEDMALCEIARNIKRKTIE 621
 QY 659 IYUITSYNNKAVEYRYKIVLEBEGKVPLEKVIYKTLSEKLEYYTEADPHVAARML 718
 D 622 ETLISDYKAKIKLVKSTVIRKLRGEPDNEBELITWAKIERLDLNNYKQLPFLVAAKALQ 681
 QY 719 AGYVSPDKIGYIVVKGGRISORAPFYMYQSDQIDVTYVYDHOIIPALRIIYGVG 778
 D 682 SGTLSKSKSGIKYIVVKGGLPNDRAEPFLVEKRRIDIEYVD-QIPRETIKLKPLG 740
 QY 779 ITEKKL-KASATGOKTLFDFLAKK 801
 D 741 VNEESLKKTNITDLDLFGASKKK 764

RESULT 6
 S35543
 DNA-directed DNA polymerase (EC 2.7.7.7) - *Pyrococcus furiosus*
 C:Species: *Pyrococcus furiosus*
 C:Date: 31-Dec-1993 #sequence_revision 02-Aug-1994 #text_change 20-Jun-2000
 C:Accession: S35543; S44596
 R:Unimol T; Ichino, Y.; Toh, H.; Aada, K.; Kato, I.
 Nucleic Acids Res 21, 259-265, 1993
 A>Title: Organization and nucleotide sequence of the DNA polymerase gene from the archae
 A:Reference number: S35543; M01D:93181200; PMID:8441634
 A:Accession: S35543

A:Molecule type: DNA
 A:Residues: 1-775 <UEM1>
 A:Cross-references: EMBL:D12983; NID:g216917; PIDN:BAA02362.1; PID:g216918
 A:Accession: S44596
 A:Molecule type: protein
 A:Residues: 1-12 <UEM2>
 A:Gene: pol
 C:Superfamily: herpesvirus DNA-directed DNA polymerase
 C:Keywords: DNA binding; nucleotidyltransferase
 F:1-775/Product: DNA-directed DNA polymerase #status experimental <MAY>

Query Match 29.54; Score 1236.5; DB 2; Length 775;
 Best Local Similarity 36.24; Pred. No. 2.3e-67;
 Matches 301; Conservative 153; Mismatches 283; Indels 95; Gaps 23;

QY 8 VLDSSYEVVNGKEPQVIIMGIAENGERVVLIDSPRPYFALLAGADPKQV---AQR 63
 D 2 ILVDYITTEGKFPVIRLF-KENKERTKHDTIRPIYALLRDSKIEVKKITEKRG 60
 QY 64 ALSRPSPIIGVEDDKRYKFGRRPRVLRITVLEAVREYELVKVNDGVEDVLEADIRF 123
 D 61 KIVR---LVDEKVEKKEKFLKPTWKVLEHPDVPFTIREKVRHRAVVDIFEDYIF 116
 QY 124 AMRYLIDHDLPEFTWYRVEAEPLNMGFRVDKYLKSRPPELYGEALAPTKLPDLRI 183
 D 117 AKRYLIDKGLI-----PMEGE-----BELKITL 138
 QY 184 APDIEVYSKQSPRPERDPIYIAVKTDDGDEVL-----FIAGKDDRKPIREVEY 235
 D 139 APDIEVLYHEGE--EFGKGPITIMISYADENBAKVIWKNIIDLPYEVVSSEREMIKFELI 197
 QY 236 VKRYDPIIDVGYNNHFMPPYLBRARILGKLDVTRRGAEPTSVAGH--VSYPGRL 292
 D 198 IREKDDPIIVYNGSPDFPLAKRAEKIGITIGRD-GSEPMQRIQDMTAVYKRI 256
 QY 293 NVDLVYAEEMPEIKISLSEVAYGVMMKSERVIINMWEIPDYDDPKKPLLOYAR 352
 D 257 HPDLVHIVIRTNIPYTLLEAVYEA--FGKREKYYAD--EIAAMSGENLERVAKY 313
 QY 353 DVPATYGLAEKIIPALQSLVYVGLPDDGVAMSVGFLEWYLIRAAFKMKELVNRVER 412
 D 314 EDKATYELGKFLPMEIQSLVGLVOPNDVRSSTGNLWVFLDKALIRNEVAPKFS 373
 QY 413 RPE-----ETRGALVLEPARGHENTIAVLDFSSMYPNIMIKYNVGPDTLVREKSG 466
 D 374 EEEYQRLRESTGSEFVEKPEKGLMENIVYLDPRALYPSIITINVASPDTLNEG---- 428
 QY 467 CGCME-APENVKRPFRCPGFFKTVLERLLELRKVRAMKYPDPDSPEYRLDEROKAL 525
 D 429 CKNYDIAIPGVGHKFCQDIPGFIPLGHLLEEROKIKTKMKE--TODIEKILLDYRQKI 487
 QY 526 KYLANASGYMGSGARWYCECAKAVTAMGRLIRTA-INTARKLQIKVIYGDPSLFLV 584
 D 488 KLANSPGYGYAYARAWYCECAESVTAMGKRYIELWKELEKPFKALYIDTIGLVA 547
 QY 585 TYD-----PEKVENFIKIKELG--FEIKLEKYKRLPFTAKKRYAGLEDGRIDI 635
 D 548 TIPGSESEIKKKALEFVKYINSLPGLLELEVEGFRGFFV-TKRRAVIDEESKVI 606
 QY 636 VGFEAVRGMCELAKEVQTKVVEIVLTSYNNKAVEYRYKIVLEBEGKVPLEKVIYK 695
 D 607 RGLIYIRRMSEIAKETQARVLEITLIKHDVEENAVRIAYEIVQCLAYEIPPEKALIV 666
 QY 696 LSKGLEEYITTEPHVAAKMLSLAGRVSPDKIGYIVVKGGRISORAPFYMYK--DP 753
 D 667 ITPPLEHYAIGPHAVAKKLAAGAVYKRGAVTIGVLRDGDGKISRA---ILAEYDP 723
 QY 754 S--QIDVTYVYDHOIIPALRIIYGVGITEKKLKASATGOKTLFDFL-ARKS 802
 D 724 KKHAYDAEYIENQVLPALVIRLIEGFGYREDLRYQKTRQVGLTSLWLNIKKS 775

RESULT 7

DNA polymerase I PAB1128 - Pyrococcus abyssi (strain Orsay)

C75023

C/Species: Pyrococcus abyssi

C/Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000

C/Accession: C75023

R/anonymous, Genoscope

submitted to the EMBL Data Library, July 1999

A/Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru

A/Reference number: A75001

A/Accession: C75023

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-771 <KAM>

A/Cross-references: GB:AJ248288; GB:AL096836; NID:95458960; PIDN:CA50625.1; PID:9545913

A/Experimental source: strain Orsay

C/Genetics:

A/Gene: polI; PAB1128

C/Superfamily: herpesvirus DNA-directed DNA polymerase

Query Match 28.8%; Score 1207.5; DB 2; Length 771;

Best Local Similarity 35.7%; Pred. No. 1.3e-65;

Matches 292; Conservative 150; Mismatches 285; Indels 91; Gaps 21;

Query 8 VLDSYEVVKEPQVLIWIGIENGERVVLIDRSFPYFALLACGADPKOY---AQRIR 63

Db 2 IILDVITBEDKPIIRIFK-KEKGEFKVEYRTFRPIYALLKDDSAIDVKKITAEKHG 60

Query 64 ALSPKSPILIGVEDDKKRYFGRPRVLRIRTVLPEAVREYELVKNVDGVEDLEADIR 123

Db 61 KIVR-----ITEVEKQKFKGRPIEWKLYLEHPQDVPAIREKIREHNAVVDIEYDIPF 116

Query 124 AMRLIDHDLFPTTWYRVEAPELENKGFVVDKYLVSREPELYGEMALPTKLPDLRIL 183

Db 117 AKRLIDKGL-----TMEGN-----BELTFL 138

Query 184 AFDEIVYKSGSPRPERRPVIVIAVKTDDGDEVL-----FLAEGKDDKRPREFEY 235

Db 139 AVDIETLYHEGE-EFGKQPIIMISYADEGAKVITWKSIDLPYEVVSEEMIKRLVKV 197

Query 236 VKRYDPDIIVGNNHDPWYLLRRARILGILDTVRVGAEPPTSVGH---VSPGR 292

Db 198 IREKDPDVIITYNGDNFDPFLYLRKAEGLIKPLGRD-NSEPKQKQKQDGLAIVEIKRI 256

Query 293 NVLDIYVAEEMPEIKISLEVAEYLGVMKSESVIINMWEIPDYWDPKRPLLOYAR 352

Db 257 HFDLPVIRRTINLPYTLLEAVEYAI-FGKSEKYYAAH--EIAEMEGKGLSEVAKSM 313

Query 353 DDVATYGLAEKILPFAIQLSVYGLPLDOYGANSVGRLEWYLIRAFKKKEIVPNRVE 412

Db 314 EDATVPELGEFFPEMEQLARLVGQPYWDVSRSTGMLVEMFLRKAYENNELAPNKP 373

Query 413 RPE-----ETRYGAILLEPLRGVHENIAVLDFSSMYPNIMIKNVGPDTLVRGEKGE 466

Db 374 EREYERLRLESYEGYVKEPEKGLMEGIVSLDFRSLYPSIITTHNVSDTLNR--ENKE 431

Query 467 CGCEHAPVVKRRCPGPFKTVLERLLELRKRVRAEMKKYPPDPSPEYRLDDEOKALK 526

Db 432 YDV--APQVGRFCKDFPGFIPSLGMLLEEROKIKKKMKK-SKDPVEKKLLDVRORAIK 488

Query 527 VLNASYGYMGSGARWYCRECAKAVTANGRLIRTAIINARKGLKLYIGDTSPLFTY 586

Db 489 ILNASYGYGYAKARWYCKECAESVTAMGROYIDLVRRELESKFKVLYIDTGLVATYI 548

Query 587 D-----PEKVENFIKIKEELG--FEIKLEKYKRLFTFEAKKRYAGLLEDGRIDVG 637

Db 549 PGAGHEIKERALKFVEYINSKLPGLLELEGFYARGFV-TKKKVALIDEBGIVTRG 607

Query 638 FEARVGMCELAKEVQTVVVEIVLKTSEVNAVEYKRVKIVELEBGKPIKLYIWKLS 697

Db 608 LEIYRRDSEIJAETQAVLVAILLKHGNVDAVKIVKEVTEKLSYELPPEKLYIYEOIT 667

Query 698 KRLSEYTTAEHVVAAKKMLSAGYRVSPGDKIGYIVVGGGRISQRAPIYMWK--DPS- 754

RESULT 8

DNA-directed DNA polymerase (EC 2.7.7.7) - Thermococcus sp.

S67920

C/Species: Thermococcus sp.

C/Date: 17-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 18-Jun-1999

C/Accession: S67920

R/Southworth, M.S.; Kong, H.; Kucera, R.B.; Ware, J.; Jannasch, H.W.; Perler, F.B.

submitted to the EMBL Data Library, January 1996

A/Description: Cloning, expression and modulation of the 3'-5' exonuclease activity of

A/Reference number: S67920

A/Accession: S67920

A/Molecule type: DNA

A/Residues: 1-775 <SOU>

A/Cross-references: GB:U47108; NID:91197451; PIDN:AAA88769.1; PID:91197452

A/Experimental source: strain 90N-7

C/Superfamily: herpesvirus DNA-directed DNA polymerase

Query Match 28.6%; Score 1197.5; DB 2; Length 775;

Best Local Similarity 36.1%; Pred. No. 5.3e-65;

Matches 301; Conservative 139; Mismatches 296; Indels 97; Gaps 24;

Query 8 VLDSYEVVKEPQVLIWIGIENGERVVLIDRSFPYFALLACGADPKOY---AQRIR 63

Db 2 IILDVITBEDKPIIRIFK-KEKGEFKVEYRTFRPIYALLKDDSAIDVKKITAEKHG 60

Query 64 ALSPKSPILIGVEDDKKRYFGRPRVLRIRTVLPEAVREYELVKNVDGVEDLEADIR 123

Db 61 TVKVKR-----AEKVQKFKGRPIEWKLYFNHPQDVPAIRDRIAPNAVVDIEYDIPF 116

Query 124 AMRLIDHDLFPTTWYRVEAPELENKGFVVDKYLVSREPELYGEMALPTKLPDLRIL 183

Db 117 AKRLIDKGLI-----PMEGD-----BELTFL 138

Query 184 AFDEIVYKSGSPRPERRPVIVIAVKTDDGDEVLFLAEGKDDKRPREFEY 233

Db 139 AFDIETLYHEGE-EFGTQPIIMISYADEGAKVITWKSIDLPYEVVSEEMIKRLVKV 195

Query 234 EYKRYDPDIIVGNNHDPWYLLRRARILGILDTVRVGAEPPTSVGH---VSPGR 290

Db 196 RYKREKDPDVIITYNGDNFDPFLYLRKAEGLIKPLGRD-GSEPKQKQKQDGLAIVEIKRI 254

Query 291 RLNVLDIYVAEEMPEIKISLEVAEYLGVMKSESVIINMWEIPDYWDPKRPLLOYAR 350

Db 255 RIHFDLPVIRRTINLPYTLLEAVEYAI-FGKPKKYYAAH--EIAQMEGKGLSEVAKSM 311

Query 351 ARDVRATYGLAEKILPFAIQLSVYGLPLDOYGANSVGRLEWYLIRAFKKKEIVPNR 410

Db 312 SMDAKTYELGEPFPEMEQLARLVGQPYWDVSRSTGMLVEMFLRKAYENNELAPNKP 371

Query 411 VERPEET-----RYGAILLEPLRGVHENIAVLDFSSMYPNIMIKNVGPDTLVRGEKGE 465

Db 372 PDREILARRGGVAGGVYKPEKGLMDNIYLDLFRSLYPSIITTHNVSDTLNR--ENKE 426

Query 466 ECGGWE---APEYKRRRCPGPFKTVLERLLELRKRVRAEMKKYPPDPSPEYRLDDEOKALK 522

Db 427 --GCKEYDVAPEYKRRCPGPFIPSLGLDGLLEBKIKRKKKA-TVDPLEKKLLDYRQ 483

Query 523 KALKVLANASYGYMGSGARWYCRECAKAVTANGRLIRTAIINARKGLKLYIGDTSPLFTY 581

Db 484 RAIKVLANASYGYGYAKARWYCKECAESVTAMGROYIDLVRRELESKFKVLYIDTGLVATYI 543

Query 582 LFVTV---DE-----KVENFIKIKEELG--FEIKLEKYKRLFTFEAKKRYAGLLEDGR 632

Db 544 LHATIPQADMEYTKKAKKELKTIINPKLPGLLELEGFYARGFV-TKKKVALIDEBGK 602

QY 633 IDIVFPAVAGDMCELAKEVQTKVVEIVLKTSEVNAKAVEYRKIVKELEBQKVPLEKVI 692
 Db 603 ITRRGLEIVRDMSEIAKESTOARLEAILKHGDVEBAVRIYKEVTEKLSKYEVPEKVI 662
 QY 693 WKLTSKRLSEYTTTAPPHVAAKRLSNGYRVSPGDKIGYIVKGGGRISORAMPYRMVD 752
 Db 663 HEQITRDLRYKATGPHVAAKRLAARGVIRPGTIVSYIVLKSSGRIGRAIPAEF-D 721
 QY 753 PS--QIDVTVYVDHQIIPALRILIGVFGITEKKLKASATGQKTLFDFLAKSK 803
 Db 722 PTKHRYDAEYVYENQVLPAYERILKAFYRKEDLRYOKTKQVGLGAWLKVKKG 774

RESULT 9

S66593

DNA-directed DNA polymerase (EC 2.7.7.7) Vent, intein containing precursor - Pyrococcus
 N:Contains: DNA endonuclease (EC 3.1.-.-) PI-1; DNA-directed DNA polymerase (EC 2.7.7.7)
 C:Species: Pyrococcus sp.
 Date: 24-Aug-1996 #sequence_revision 01-Nov-1996 #text_change 24-Sep-1999

Accession: S66593
 Name: M.Q.; Southworth, M.W.; Merzha, F.B.; Hornstra, L.J.; Perler, F.B.
 Submitted to the EMBL Data Library, August 1993
 A:Description: In vitro protein splicing of purified precursor and the identification of
 A:Reference number: S68593
 A:Accession: S68593
 A:Molecule type: DNA
 A:Residues: 1-1312 <XUA>
 A:Cross-references: EMBL:U00707; NID:9436492; PID:AAA67130.1; PID:9825735
 R:Xu, M.Q.; Southworth, M.W.; Merzha, F.B.; Hornstra, L.J.; Perler, F.B.
 Cell 75, 1371-1377, 1993
 A:Title: In vitro protein splicing of purified precursor and the identification of a bra
 A:Reference number: S68581; MUID:94094330; PMID:8269515
 A:Contents: annotation
 R:Xu, M.Q.; Comb, D.G.; Paulus, H.; Noren, C.J.; Shao, Y.; Perler, F.B.
 EMBO J. 13, 5517-5522, 1994
 A:Title: Protein splicing: an analysis of the branched intermediate and its resolution b
 A:Reference number: S52065; MUID:95080235; PMID:7988548
 A:Contents: annotation; self-splicing mechanism
 C:Function: <NUCL>
 A:Description: nucleotidyltransferase
 A:Note: DNA-directed DNA polymerase Vent
 C:Function: <ENDO>
 A:Description: endonuclease; hydrolase
 A:Note: DNA endonuclease PI-pepi
 C:Superfamily: hypothetical protein PH0202
 C:Keywords: DNA replication; endonuclease; hydrolase; nucleotidyltransferase; protein sp
 1-491,1030-1312/Product: DNA-directed DNA polymerase Vent #status predicted <MT1>
 492-1039/Product: DNA endonuclease PI-1 (pol Vent intein 1) #status predicted <MAT2>
 F/1030-1312/Domain: DNA-directed DNA polymerase Vent extein 2 #status predicted <XT2>
 F/491-1030/Cross-link: peptide (Ala-Ser) #status experimental

Query Match

23.1%; Score 965.5; DB 2; Length 1312;

Best Local Similarity 21.9%; Pred. No. 1.4e-50;

Matches 300; Conservative 154; Mismatches 282; Indels 635; Gaps 23;

QY 8 VLDSYEVVKEPVIYIINGIENGERVYLIDSFPFYFALLAGADPKOV-----AQRIR 63
 Db 2 IIDADYITDEGKPIRIFK-KENGEFVEYDNRFRYIYALLKXDSQIDVRKTIERFG 60
 QY 64 ALSRPSPITIGVEDKRYKRGPRVRLIRTVLPEAKREYELVKNADGVEDLEADIRF 123
 Db 61 KIRK-----IIDAEKRRKFFLDRIPLVWRKLFEHPQVPAIRDKIRSHSAVIDIFEDIRF 116
 QY 124 ANRYLIDHDLFPFTWYRVEAEPLNKKGRFVDDKVIVLSRPELYGEALAPTKLPDLRL 183
 Db 117 AKRYLIDKGLI-----PHEGD-----EELKLL 138
 QY 184 AFQIEVYSKQSPREBRDPVIVAVKTDGDEVL-----FLAEKDDRKPIRFVEY 235
 Db 139 AFQIEVLYHEGE-EPAKGPITIMISYADEBAKVIWTKKIDLPYVEVSSREMIKRFV 197

QY 236 VKRVDPIIIVGNNHFMPIYLLRRARILGILKDVTRVGAEPPTSVHGH---SVVGRSL 292
 Db 198 IREKDDPVIITNGSDPDLVYRAEKGIKILPGRD-GSEPKQORLIGMTVVEIKGR 256
 QY 293 NVLDYVAEMPEIKIKSLSEVAEYLVGMKSERVIMMWEIIPYWDPKRRPLLOQVIR 352
 Db 257 HPDLVHVRITINLPYTLFVAYEAI-FGKPEKVVAAH-EIAEMETGGLERVAKYM 313
 QY 353 DVPRATYGLAEKILPFAVLSYVGLPLDOVGAMSGVRFLEMYLIRAFKQKELVNRVE 412
 Db 314 EDAKVYTELGRFFPMQAQSRVLVGOPLMDVSSSTGNLVEMYLILRAKARNEILANKPD 373
 QY 413 RPE-----ETRYGALVEPLRGVHEIADVLPSSMYPNIMIKYVGPDTLVPRSGE 466
 Db 374 EREYERLAEISYAGVYKEPEKLMGQIVSLDRSLYSIIITINVSPTLNE----- 427
 QY 467 CGGME---APYKRRFRKCPGPFKTVYERLLEAKKRAEMKYPDPSPEYRLDEROK 523
 Db 428 -GREDYVAPSEVGHKCFKDFPLSLKRLIDERSQETIKRMKA-SKDPLEKQMLDYROR 485
 QY 524 ALKYLVA----- 529
 Db 486 AIKILANSILPEBWPPLIKNGKVIIRIGDPVDGLMKANOKKXKGTDELEVAGIHA 545
 QY 530 ----- 529
 Db 546 SPDRSKKARVAVAVAIRHRSGNVYRIVANSGRKITTEGHSLFYVRNGDLVEATGED 605
 QY 530 ----- 529
 Db 606 VKIGDLAVPRGVLNPEKERNIVLELLNLSPEETEDILITPYKGRKNFPGMLRTL 665
 QY 530 ----- 529
 Db 666 WIRGEKRVRTASRYLRLEMLGIRLAKIGYDIIDKSGLEKRYTVLEKLVAVRYNNGK 725
 QY 530 ----- 529
 Db 726 RELVFEFNAVVRVILMEPEELKEMRLGTRNGFRNGTVIDEDPAKLLGYVSEGSARK 785
 QY 530 ----- 529
 Db 786 WKQGTGMSYVRLYNENDEVLDMEHLAKKFFGKVKGKNYVEIPKKMAYIIFESL 845
 QY 530 ----- 529
 Db 846 LAENKEVPEVITSSKGVNAFLGEGYFIDGQVHPSKVRSLSTKSELLVNLINSIG 905
 QY 530 ----- 529
 Db 906 VSAIKLQDSGVYRVYVNEELKFTERYKKKNVYSHIVKDIKETFQKVFQKNISYKFF 965
 QY 530 ----- 529
 Db 966 RELVENGKLDREKAKRIEMLNGDIYLDRIYELKEBYDGVYVOLDSDDEBNFLAGFRL 1025
 QY 530 ---NAGYTMGWSGARWYCRECAKATYAMGRHLIR-TALNIRKLGLKVIYGTDSLFT 585
 Db 1026 YANNSYVGYGAKARWYCRECAESYAMKREYIEYVKELEBKFGRKLYLITDGLYAT 1085
 QY 586 Y---DEPKY---NFIKIYIEELG--FEIKLEKYTKLFTFEAKKRYAGLEDGMDIV 636
 Db 1086 IPGAKPEELIKKALFEVYVINAKLPGLELELEGVGVGFV-TRKKYALIDEGKILIR 1144
 QY 637 GFPAVAGDMCELAKEVQTKVVEIVLKTSEVNAKAVEYRKIVKELEBQKVPLEKVI 696
 Db 1145 GLEIVRDMSEIAKESTOAKVLEAILKHGVEEBAVRIYKEVTEKLSKYEVPEKVI 1204
 QY 697 SKLSEYTTTAPPHVAAKRLSAGYRVSPGDKIGYIVKGGGRISORAMPYRMVD----- 752
 Db 1205 TRPLHEYKAIIPHVAAKRLAARGVIRPGMIVIGYIVLRGDPISKKA---ILAEFDR 1261
 QY 753 PSQIDVTVYVDHQIIPALRILIGVFGITEKKLKASATGQKTLFDFLAKSK 803

Db 1262 KHKYDAEYIENOVLPALVILBAFGYKEDLBMQKTKQTGLTAMLNKKK 1312

RESULT 10

C71210

probable DNA-directed DNA polymerase - Pyrococcus horikoshii

C/Species: Pyrococcus horikoshii

C/Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 21-Jul-2000

C/Accession: C71210

R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin

M.; Ohtuku, Y.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; Kushida, N.; Oguchi

DNA Res. 5, 55-76, 1998

A/Title: Complete sequence and gene organization of the genome of a hyper-thermophilic

A/Reference number: A71000; MUID:98344137; PMID:9679194

A/Accession: C71210

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-1235 <RAW>

A/Cross-references: GB:AP000007; NID:93236134; PID:BA031074.1; PID:93258391

Experimental source: strain OT3

Note: this accession replaces an interim accession for a sequence replaced by GenBank

C/Genetics:

C/Superfamily: hypothetical protein PH0202

Query Match

Best Local Similarity 23.8%; Score 955.5; DB 2; Length 1235;

Matches 299; Conservative 145; Mismatches 291; Indels 559; Gaps 24;

QY 8 VLDSSEYVWKEPQVLIWGIAENGERVVLIDRSPPRYFALL---APGADPKVAQRIR 63

Db 2 ILDDADYITDEDKPIIRIK-KENGEFKVEYBNRRPYIYALLRDSALDEIKITTAQHG 60

QY 64 ALSPEKPIEDVEDDKKIFGPRPRVIRIYVLPVAVREYELVKNVGVDEVDLADIR 123

Db 61 KVVV---IVETEKIQRFKFLPEVWKLYLEHPQDVPAIRDKIRHRAVVDIFEYDIPF 116

QY 124 AMRLIDHDLPRFTMYRVEAPLENKGFRRDQVLYVSRFEPLGALAPKLPDLRL 183

Db 117 AKRYLIDKGL---TPMEGN-----EKLTF 138

QY 184 AFDEIVSVKSGSPRPDPVIVIAVKTDDGDEL-----FLAEGKDRKPIREFVEY 235

Db 139 AVDIETLYHEBE-ERKGPVIMISYADEGAKVITWKIKIDPIYEVSVSEEMIKRLIRV 197

QY 236 VKRYDPDIIVGYNNHDPYLLRRARILIGLDVTRVGAEPITTSVGH---VSPGRL 292

Db 198 IKERDPRVITYNGDNDFPYLLKRAEKLGIKL-LAGRDSEPKMKQKGDLSLAVEIKRI 256

QY 293 NVLDYVAEMPEIKISLEVAELGVKSKSERVITIMWEIPDYMDPKPKPLLOQYAR 352

Db 257 HFDLPVIRRTINLPYTLAEVYEI-EGKPEKRYAD--EIAKAMETGEGELERAKYSM 313

QY 353 DDVATYGLAKIPLFAIQLSVYTGLPIDQVAMSVGRLEWYLIRAFKXKELVNVVE 412

Db 314 EDATYTYLGEFPEFMEQLARLVQDPWDYSSRSTGLVEMFLRKRYKERNELAPNPD 373

QY 413 RPE-----ETRYGAILPEPLRGVHENIAVLDFSSMYNIMIKVNGPDLTPREKCGE 466

Db 374 EKEVERRLRESYEGGYVEPEKGLWEGIVSLDFRLYSIIITHNVSDTLNRE----- 427

QY 467 CGGHE--APVGRFRRCPPGFCTVLERLLLEAKVRAEMKATPPSPRYLIDEXOK 523

Db 428 -GCEBYDVAIVGRHFCDFPGFIPSLIGGLEEROKIKKQMK- SKDPVEKGLDYQR 485

QY 524 ALKYLVA----- 529

Db 486 AIKILANSILDEMLPIYENKVRVKTGDPIDREIENARVYKADGETILLEVOKLAL 545

QY 530 ----- 529

Db 546 SPNETKSELKVVALLRHRYSGKVAISIKLSGRRIKITSGLIFSVANGKLAVKRGDE 605

QY 530 ----- 529

Db 606 LKPDLLVVBGRLLKLPESKQVNLVELLLKLPBEETSNIWMMI PVKGRNPFKMLKTLV 665

QY 530 ----- 529

Db 666 WIRGEGRPRTAGRYLKHLERLGVYVKKRGCEVLDMESLRYKLYETLIKRLKTNNS 725

QY 530 ----- 529

Db 726 RAVVVEFNSLRDVSLMPILBELKEMIIGEBRPGKIGFIDVDDSPAKLGYISGDEK 785

QY 530 ----- 529

Db 786 DRVFKSKQNVLEDIAKLAEKLGKVRGRGYEVSGKISHAIFRYLAEKRIPEFIPT 845

QY 530 ----- 529

Db 846 SPMDIKAFILKLGNAEELTSTKSELVNLILLNSIGVSDIKIEHKGYRYRYINK 905

QY 530 ----- 529

Db 906 KESNGDVLDSVESIEVEKYGWYVLDSEDNENFLVGRGLYARNYGYGYAKAR 965

QY 544 YCECAKAAVYANGRHILRTINARLKLKLYIGDTSLEFVY---DPEKVE---NFI 595

Db 966 YCECAKAAVYANGRHILRTINARLKLKLYIGDTSLEFVY---DPEKVE---NFI 595

QY 596 KIKEELG--FEIKLEKVVYRLPFTEAKRRYAGLLEDEGRIDYGFVAVRDMOELKEVQ 653

Db 1026 DYNSKLPGLYLEEYEFYARGFV-TKKRYALIDEGKLVYTGLELYVRDMSEIATETQ 1084

QY 654 TKVVEILKTSSEVNAKAVYRKIVKLEBQVPIEKLVITKLSKLEEYTTAPHVAA 713

Db 1085 ARVLEAILKKGAVNAEKIVQDVTEKLTNYEVPEKLVYEQTRIPNEKALGPVAVA 1144

QY 714 KRLMSAGYRSPQDKIGYVYVGGGRISORAMPYMWK--DP--SQIDVYVYDHQIIPA 769

Db 1145 KRLMARIGIKVPGWIVIGYIYLRDGPISKRA---ISLEEDPRKHKVDAYIENOVLP 1201

QY 770 ALRILGYGITEKKLKASATGOKTLPDFL-AKKS 802

Db 1202 VERILKAFYKREDLRMQTKOVGLAMIKVKS 1235

RESULT 11

C69028

DNA-dependent DNA polymerase family B (PolB) - Methanobacterium thermoautotrophicum (b

C/Species: Methanobacterium thermoautotrophicum

C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999

C/Accession: C69028

R:Smith, D.R.; Doucette-Stamm, L.A.; DeJonghe, C.; Lee, H.; Dubois, J.; Aldredge, T.;

J. Qu, D.; Spadefora, R.; Viciore, R.; Wang, Y.; Mierzowski, J.; Gibson, R.; Jiwani, N

K.; Smith, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.

J. Bacteriol. 179, 7135-7155, 1997

A/Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: func

A/Reference number: A690007; MUID:98037514; PMID:9371463

A/Accession: C69028

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1586 <MTH>

A/Cross-references: GB:AE000888; GB:AE000666; NID:92622304; PID:BA085697.1; PID:9262231

A/Experimental source: strain Delta H

C/Genetics:

A/Genetics: MTH1208

Query Match

Best Local Similarity 21.8%; Score 915; DB 2; Length 586;

Matches 218; Conservative 110; Mismatches 226; Indels 62; Gaps 12;

QY 1 MTEVVFVLDSSYEVWKEPQVLIWGIAENG--ERVVLIDRSPPRYFALLAPGADPKOV 58

Db 1 MEDYRNVLIDIDYTVDEVPVIRLFGKSGNGNEPIAHDRSFRPVIYAI-----PTDL 54

Qy 59 AGRIRALSRRKSPILIGVEDDKRYKGFPRPVLRIRTVLPEAREYRELYKRVQGVDE 118

Db 55 DECLRELELEKLEKLEKLEKLEKLEKLEKLEKLEKLEKLEKLEKLEKLEKLEKLE 112

Qy 119 ADIRFAMRYLIDHDLFP-----FTWYVEAEPLKNGRVDKYLKSRPEELXGELA 173

Db 113 HDIPFRRYRLIDKSLVPMHELEFGQVVDAPSVTTDRTVTEVYNGRQSGSANG----- 168

Qy 174 PTLPLRLIADIEVSKOSPPREPDPIYVAVK-----TDDGDEVLPIAEGD 224

Db 169 -----LDILSPDIEVRKPMDEKDELMIGVANGMGEVSISTRGHDHDEVEVED 222

Qy 225 DRKPIREPEYKRYDPIIYGVNNHEDMFLRRARILGIKLDV-----TRVGA 276

Db 223 ERRLERFAEIVIDKKPDLVGNDSNDFDPYTRRPAALIGAEIDLGDGSKIRTRRGF 282

Qy 277 EPTTVHGHVSVGRNLVDYAEEMPEIKISLEEVAYLVGWKSERVIINMEIIPD 336

Db 283 ANATAIKGTV-----HVDLYPMRRYMLDRYTLERVOELFGEKIDLPGRIME--- 333

Qy 337 YMDPKRPLLIQVARDVRYGLAEKILPFAIQSVYVGLPLDQVGMASVGRLEWTL 396

Db 334 YMDRELRLDELFRYSLDDVAVTRIAEKILPLNLRLVGLPDLISRMATGQAEWFL 393

Qy 397 IRAAFKMKELVNRVERPEETFR-----GATVLEPLRGVHENTAVLDPSGMYPNIMIKY 450

Db 394 VRKAYGELVKNKPSQSDFSRRGRAGVGYKEPEKLAHNTIVQDFRSLVPSIILSK 453

Qy 451 NVGPDPLVARGKCGCGCEGMEAPVGRFRCPGPFKTVLERLLERKVRAMKPYR 510

Db 454 NISPDILTDEES-----ECYAPPEYGRFKSPRGVPSVISELSEKVRKKEBKG-SD 508

Qy 511 DSEYRLDEROKALKVLNANSGYMGSGARWYCECKAKAVTANGRLIRTAINTARL 570

Db 509 DPMERKILNVQOELAKRLANMVGYSFRWYSMECKEALITAMGRDYIKKITKTAEEF 568

Qy 571 GLKVIYGDTSLEFVY 586

Db 569 GFHTVYADTDGFYATY 584

RESULT 12

T40242

DNA polymerase delta large chain - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000

C:Accession: T40242

Bozzym, K.; Beck, A.; Reinhardt, R.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G. submitted to the EMBL Data Library, October 1999

A:Reference number: 221916

A:Accession: T40242

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Release: 1-1086 <BOR>

A:Cross-references: EMBL:AL121815; PTDN:CAE58156.1; GSPDB:GNO0067; SPDB:SPBC336.04

A:Experimental source: strain 972h; cosmid 6336

C:Genetics:

A:Gene: SPDB:SPBC336.04

A:Map position: 2

A:Insertions: 77/1

C:Superfamily: herpesvirus DNA-directed DNA polymerase

Query Match 18.4%; Score 771.5; DB 2; Length 1086;

Best Local Similarity 27.3%; Pred. No. 6,7e-39;

Matches 246; Conservative 160; Mismatches 319; Indels 177; Gaps 30;

Qy 3 EVVFTVLDSSYEVGKEPOVILKGIANGSRVYLIDRSRPYVALLAPGADPKQVAGKI 62

Db 97 DIFQGIIDSEFTBSVSIRLREVTNGNSILVHVGLPIFYKAPVGRPEMLKRT 156

Qy 63 RAA-----SRKSPILIGVEDDKRYKGFPRPVLRIRTVLPE 98

Db 157 QDLDATONGGVLDHCIEEMKENLYFGQNGKSPFIKI-----FTNPRILSRARVFR 210

Qy 99 AVNEVELVANGVDEVDLEADIRFAMRYLIDHDLFPFTWYVEAEPLKNGRVDKYLKSRPEELXGELA 154

Db 211 GERNFELFPVGVG-TTFESNTQYLRIMIDDDVGNMGMHILASKQGRYONRNSQ 269

Qy 155 -----DKYLVRSRBEPLXGELAFTLPLRLIADIEVSKOS-SPREPRPVY 205

Db 270 IEMINIKDLISL-----PAGGW--STGAFLRMSDLECGRKRVPDSIDPVLQ 320

Qy 206 IAVKTDGDEVLPIA-----EGKDRKPIREFVEYKRYDPIIYGV 247

Db 321 IASIVQYDSTPFRVNVFVCDTCSQIVGTQVYEFQNAEMLSWSKRVADVPDVLIGY 380

Qy 248 NNNHFPMPYLLRRARILGI-----KLDVTRRVGAEPPTSVHGH-----VSVGRUN 293

Db 381 NICNFDIPYLLDRAKSLRINHFPPLIGRIHNFPSVAKETTFSSKAYGRRESKTSIGRLQ 440

Qy 294 VDLVDAEEMPEIKISLEEV-AYELGWMKSERVIINMEIIPYMD--DPKKRPLILQ 349

Db 441 LDMQVWQDRDPLRSYSINAVCSQFLGQEKED---VHSIITDLOGTADSRRL--LAI 494

Qy 350 YARDVRYATYGLAEKILPFA--IQSVYVGLPLDQVGMASVGRLEWTLIRAFKMKELV 407

Db 495 YCLKDAVLPQRLMDLGMCFVNYTEMARVCPFNPLARGGQIKVIGLFRKXLDHDLV 554

Qy 408 PN-RVERPEETFRGATVLEPLRGVHEN-IAVLDPSGMYPNIM----- 447

Db 555 PNIRVNGTDEQEGATVLEPIKGYTDPILITLDSLSISIQAHNLCYTTLLDSTAE 614

Qy 448 -----IKNVGP--DLYVRGKCGCGCEGMEAPVGRFRCPGPFKTVLERLLERK 499

Db 615 LKAKODVDSVTPNDGYFVKPHVR-----KGLPILTLADLNAAR 654

Qy 500 RVRAMKPYRPPSPRYRLDEROKALKVLNANSGYMGSGARWYCECKAKAVTANGRL 559

Db 655 KAKADLKK-ETDPFKAVALDGRQALAKVANSYVGTGATNGRPLCALISSVSYGRQM 713

Qy 560 IRTAINIARL-----GLKVIYGDTSLEFVYDPEKVENPFIKIKELEF----- 604

Db 714 IEKTVVEKRYRIENGYSHDAVITYGDTDSVAVKGVYKTLPEAMKLGEEAAYVSDQFP 773

Qy 605 ---EIKLEKYYRLPFEAKKRYVAGLL-----EDGRIIDVGEAARGMCELAKEVQTVK 656

Db 774 NPIKLEBEKPYR-YLISKSKRYAGLPWTRTDVYDMOSKIGETVYRDNCPISVYIDRA 832

Qy 657 VEIVLKTSSVNRNAVYRKIVKELEBGKPIELKLVTKTSLGRLEETAPVYVAKEM 716

Db 833 LRKRLIDQVBOGOLFTYKAVISDLQNKIDMSQVITTKALSK--TDVAAKAAVELEARN 890

Qy 717 --LSAGRRSPGDKIGYIVVG--GGRISQRAV-PYRWKQDSQIDVTYVHQIIPAL 771

Db 891 RKRDGSAPIADIRVAVYIIRGAQGDQFYMRSEDPYVLENNIPIDAKYILENQLSKPIL 950

Qy 772 RI 773

Db 951 RI 952

RESULT 13

S71551

DNA-directed DNA polymerase (EC 2.7.7.7) KOD, intein containing precursor - Pyrococcus

N:Contains: DNA endonuclease (EC 3.1.1.-) PI-1; DNA endonuclease (EC 3.1.1.-) PI-11; DN.

C:Species: Pyrococcus sp.

A:Variety: strain KOD1

C:Date: 22-Nov-1996 #sequence_revision 22-Nov-1996 #text_change 20-Apr-2000

C:Accession: S71551

R:Kakihara, H.; Takagi, M.; Imanaka, T.

submitted to the EMBL Data Library, March 1994

A:Description: Cloning and analysis of the DNA polymerase gene from a new hyperthermophilic

A:Reference number: S71551

A:Accession: S71551

A:Molecule type: DNA
 A:Residues: 1-1670 <KAK>
 A:Cross-references: EMBL:D29671
 C:Function: <NUC1>
 A:Description: as DNA-directed DNA polymerase, catalyzes the polymerization of DNA at th
 A>Note: DNA-directed DNA polymerase KOD
 C:Function: <EN1>
 A:Description: as DNA endonuclease PI-I, catalyzes the hydrolysis of internal phosphodi
 A>Note: DNA endonuclease PI-PspI
 C:Function: <EN2>
 A:Description: as DNA endonuclease PI-II, catalyzes the hydrolysis of internal phosphodi
 A>Note: DNA endonuclease PI-PspII
 C:Superfamily: DNA-directed DNA polymerase KOD
 C:Keywords: endonuclease; hydrolase; nucleosidyltransferase; protein splicing
 F:1-406/767-851/1388-1670/Product: DNA-directed DNA polymerase KOD #status predicted <NM
 F:1-406/Domain: DNA-directed DNA polymerase KOD extein 1 #status predicted <XT1>
 F:407-766/Product: DNA endonuclease PI-I (pol KOD extein 1) #status predicted <MAT2>
 F:767-851/Domain: DNA-directed DNA polymerase KOD extein 2 #status predicted <XT2>
 F:852-1387/Product: DNA endonuclease PI-II (pol KOD extein 2) #status predicted <MAT3>
 F:1388-1670/Domain: DNA-directed DNA polymerase KOD extein 3 #status predicted <XT3>
 F:406-767/Cross-link: peptide (Arg-Ser) #status predicted
 F:851-1388/Cross-link: peptide (Asn-Ser) #status predicted

Query Match 18.4%; Score 768.5; DB 2; Length 1670;
 Best Local Similarity 17.5%; Pred. No. 1.8e-38;
 Matches 302; Conservative 146; Mismatches 286; Indels 989; Gaps 24;

QY 8 VLDSSEYVNGEPVYIIGLWNGERVLIRSPRYVYALLAGADPKQV---AQRIR 63
 DB 2 ILDDYDYTEDEKPYIRIFK-KENGEFKLEYRTEPEYFALLKDSALIEYVKITAEHNG 60
 QY 64 ALSRPKPIIGVEDDKRYKFRPRPVLRITVLPVAVSEYELVNVGVNVDVLEADIRF 123
 DB 61 TVTVVYK---VEKQKQKFLGRPVVKLYFTYHPDPAIKDKIREHGAVIDIYBIDPF 116
 QY 124 AMRYLIDHDLPTFTWYRVEAPLENGKGRVYDKVYLVKSRPEPLYGALAPDKL 183
 DB 117 AKRYLIDKGLV-----PHEGD-----BELKML 138
 QY 184 AFDIEVYSKQSPREPRPVIVIAKTTDDGEVL-----FLAEGKDKRPREFVEY 235
 DB 139 AFDIQTLYHEGEERAE-GRIMISYADEGARVITWKNVDLPYDVVSTEREMIKRFLRV 197
 QY 236 VKRDPDPIIVGNNHFPWYLLRRARILGKLDVTRVGAAPTSSVNGH---YSVGRQL 292
 DB 198 VKEDPDLVLTNGDNDFAYLKRCCEKLGINPALGRD-GSEPKIQKMGDFAVEVKRI 256
 QY 293 NVLDYDAEEMPEIKISLEAEVAYLGVKKSERVITINMEIPYWDPPKRPILLQYAR 352
 DB 257 HFDLYPVIRRTINLPYTLAEVVEAVFQGR-KKYVAE--EITPAMEGEMLEKVARISM 313
 QY 353 DDVATYGLAKKILPFAIQLSYVTGLPLDQVAMSVGRLEWYLLRAAFKKEILVNRVE 412
 DB 314 EDACVYELGEPFPMENQSLRLIGOSIMDVSRSGNLVEMFLRKAYEBNELAPNKP 373
 QY 413 RPE---ETRYGALVLEPLRGVHENTAVLP----- 439
 DB 374 EKELARRQSGTEGGYVKEPERGEMENIVYLDPRCHPADTKVVKGGIINISEVQGDYV 433
 QY 440 ----- 439
 DB 434 LGIDGQVRKVMEDYKGLVNINGLKTPTNHLKPVVTKNERQTRIRDSLAKSFLTKKV 493
 QY 440 ----- 439
 DB 494 KGKILTPPLFYEIGRATSENIPEBEVLKGLAGILLAEGLLRKDVVEYDSRRKRISH 553
 QY 440 ----- 439
 DB 554 QYRVETITGKDEEPRDRITTYIFERLFGITPSISEKGTNAVTLKAKQNVYLKVEIMD 613
 QY 440 ----- 439

DB 614 NIESLHAPVLGRGFEEEDGSVNRVRISVATQCTKNEMKIKLVSKLSQLGIPHQYTYQ 673
 QY 440 ----- 439
 DB 674 YOENKDRSHYILEITGKDLILFQTLIGFISERKNALNKALISOREMNLNENGFYLS 723
 QY 440 ----- 439
 DB 734 EPNVSTYEGKQVYDLTLESTPYFANGILTHNSLYPSIIITHNVSPTDINRE----- 786
 QY 468 GCME---APEVKAHFRRCPPGFETVLERLLELRKVRALKKKPPSPRYRLDEROKA 524
 DB 787 GCKEYDVAPOVGRFCDFGFIPLSLGLDLLEEROKIKKKKA-TIDPIRKLIDYQRA 845
 QY 525 LKVLV----- 529
 DB 846 IKILANSILPEEWLPVLEGEVHFVRIGELIDRMMEENAKKVRBGEVLEVSGLVPS 905
 QY 530 ----- 529
 DB 906 FNRRTNAELKRYKALIRHDYSGKVYTIIRLKSGRIRIKITSGHSLFSVNGELVEVTDEL 965
 QY 530 ----- 529
 DB 966 KPGDLVAVPRRLBELPERNHVNLVLELLGTPEBETLDIWTITPVKKKQKFFKMLRTLR 1025
 QY 530 ----- 529
 DB 1026 IFGEKRPRTARRYLRLLEDLGYVRLKKIGYVLDMSLKNRYRLVAVENRYNNGNR 1085
 QY 530 ----- 529
 DB 1086 EYLVFNISIRDAVGIMPLKELKEMKIGTLNGFMRKRLIEVDESIALKLGYVSEGYARKQ 1145
 QY 530 ----- 529
 DB 1146 RNPKNQSYSVKLVNEDPEVLDMEKLASFPGKVRGRNRYVEIPKIGYLLPEMCGVL 1205
 QY 530 ----- 529
 DB 1206 AENKRIPEVFTSPKGVRLAFLGYSAMATSTEQETQALNEKRALANQLVLLNSGVGS 1265
 QY 530 ----- 529
 DB 1266 AVKLGDGVYRYVINEELPVLKDKKQAYSHVLPKEVLSEVFGVKQKNSPQTFRK 1325
 QY 530 ----- 529
 DB 1326 MWEDGRLDPEKQRLSWLIGDVVLDREVSVVDYEDIGYVYDLSVEDNENFLVFGLVYA 1385
 QY 530 -NMSYGMGSGARWYCRECAKAVTAMGRHLIRTAI-NIARKLGLKVIYGDTSLEFVY- 586
 DB 1386 HNSYGYGYGARARWYCKEACASVTAMGREYIMTIKEIEKGFVIYDGTGFAFITP 1445
 QY 587 -DPEKVE---NPIKIKEL--GFEIKLEKYKRLFPEAKRYAGLIEDGRIDIVGF 638
 DB 1446 GAALAEYKKKAMEFLVYNKALPGALELEVEGYKGFV-TKKYAVVIDEBSKITRGL 1504
 QY 639 EAVRGDWCELAKEVQTKVAVILKTSVNVKAVYVRIKIVALEBGRVPEIKVIYMTLSK 638
 DB 1505 EIVRRDMSEIAKETQARVLEALKDGDVEKAVRIKVEIKLSKYEVEPEKLVIHQDITR 1564
 QY 699 RLEBYTTEAPHVAARMLSAGYRVSPGDKIGYIVYKGGIRISQAMPYPMVNDPS--QI 756
 DB 1565 DLKDYKATGPHVAVAKRLARWVKIRPGVITSYIVLKGSSRIDRAIRP-FDEEDPIKHKY 1623
 QY 757 DVTYVYDHOIIPAAKILGFTTEKKLKASATGQTLDFPLAK 801
 DB 1624 DAEYIENQVLPAVERILRAFGYRKEDLRQKTRQVGLSAMLKPK 1668

RESULT 14
 T43266

DNA-directed DNA polymerase (BC 2.7.7.7) delta chain - fission yeast (*Schizosaccharomyces*
N. Alternate names: DNA polymerase delta
C. Species: Schizosaccharomyces pombe
C. Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 21-Jul-2000
C. Accession: J43266
R. Biomed. G. Bouvier, D. de Recondo, A.M.; Baladacci, G.
J. Mol. Biol. 222, 209-216, 1991
A. Title: Characterization of the POL3 gene product from Schizosaccharomyces pombe indica
A. Reference number: S19661; PMID:92071954; PMID:1960723
A. Accession: J43266
A. Status: Preliminary; translated from GB/EMBL/DBJ
A. Molecule type: DNA
A. Residues: 1-1086 <PTG>
A. Cross-references: EMBL:J07734; NID:J913383; PIDN:AAA5303.1; PID:J9173384
C. Gene: pol3
A. Introns: 77/1
C. Superfamily: herpesvirus DNA-directed DNA polymerase
C. Keywords: DNA binding; DNA replication; nucleotidyltransferase

Query Match 18.2%; Score 763.5; DB 2; Length 1086;
 Best Local Similarity 27.2%; Pred. No. 2e-38;
 Matches 245; Conservative 160; Mismatches 320; Indels 177; Gaps 30;

3 EVVFTVDSSEYVGVKPEVITIGIAENGERVYLIDRSFRYPYFALLARQADPKVOAQR 62
 97 DIVFOIDSEFTGSGVPSIRLFGVINDGNSITLVVGFDPYVYKAVGFRPBEMLRFT 156
 63 RAL-----SRKSPITGVDDKRYKFGRRVLRITVLRPE 98
 157 QULDATKCGVITHCITEMENKXGQGNESPFKLT-----FTTNRLSLARKNVEER 210
 99 AVREKRELVAQVGVDEVDLEADIRPMRYLIDHDLFPFTYRYAELEPNKGFVY--- 154
 211 GFNFPEELFPVGVGV-TTFESNTOYLRFMIDCDVGNMTHLPASKYQFRQNRSMCO 269
 155 -----DKVYLVKSRPELVGEALAPTKLPDLRIILAFDIEVYSKQG-SPRPRDPYIV 205
 270 IEAMINVKDLISL-----PAGCQW---SKNAPLRIMSFDECGRGRGPPSPSIDPYIQ 320
 206 IA-VKTDGDEVLEFLA-----EGDKRKRIFRFEVYVRYKVPDIIYGV 247
 321 IASIVTQYGDSTPFRVNVFCVDTQSQIVGTQYVEFQNOAEMLSSKFRVADVDPIVLGY 380
 248 NNNHPPMPYLLRRARIIGI-----KLIVTRVGAEPPTSVHGH-----VSVGRIN 293
 381 NICNFDIPYLLDRAKSLRINPFLGRINHPFSVAKETTPSSKAYGTRRESKTSIPGRLO 440
 294 VDLVYAEEMPEIKIKSLLEV-AEYGVWKKSERVIIMMWIPYMH---DPKGRPLLO 349
 441 LDMQVQMRDPLKLSYSLANVCSQIFGEQKED---VHISITIDLVNGTDSRRK--LAI 494
 350 YARDVRYATYGLAKIIPFA--IQLSYVTGLPLDVGVAMSGVFLIEMYLIRAPFKKELY 407
 495 YCLKQAYLPORLMDKMLCFVNTYEMKRVGVPEFNLLRGGQIVVLSQFLCKALQHDLVV 554
 408 EN-RVERPEETYGAIYLEPLRGVHEN-IAVLDSSSYPMN-----KGLLPILLADLLNARK 447
 555 PIRVNGTDEQYEGATVIBIKGYDTPATLDPSLSIMQAHNLCTYTLIDNSTAEL 614
 448 -----IKYVNP--DTLVAPGKCECCGCEAPVKARFRFRCPGFFKTVLERLLEAK 499
 615 LKLKQDVYSVTPNGDYFVFKPHR-----KGLLPILLADLLNARK 654
 500 RVRAEMKKYPPDSPERYLLDEROKALKVLNANYSYGMGSGARVYCECKAKATVANGRL 559
 655 KKKADLKK-ETDPFKAVLDGRQLAKVANSVGFATATNGRLPCALISSSTVSGRQM 713
 560 IETAINIRKL-----GLKVIYGDPSLFYTYDEKENVIKIKELGFI-----604
 714 IKTQVVEKRYRIENGYSHDAVYIIGDITDSYVAFVGVKTIPEMKKLGEEAANYSVDQRP 773
 605 --ETIKERVYKRLFTPEAKGRVAGLL-----EDGRIDIVGFEAVRGDMCELAKEVQTKV 656

DNA-directed DNA polymerase (BC 2.7.7.7) delta catalytic chain - human
A. Species: Homo sapiens (man)
C. Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C. Accession: A41618; S35455
R. Chung, D.W.; Zhang, J.; Tan, C.K.; Davis, E.W.; So, A.G.; Downey, K.M.
Proc. Natl. Acad. Sci. U.S.A. 88, 11197-11201, 1991
A. Title: Primary structure of the catalytic subunit of human DNA polymerase delta and ci
A. Reference number: S35455; PMID:92178967; PMID:1542570
A. Accession: A41618
A. Status: Preliminary
A. Molecule type: mRNA
A. Residues: 1-1107 <CHU>
A. Cross-references: GB:M0397; NID:J9181619; PIDN:AAA5439.1; PID:J9181620
R. Yang, C.L.; Chang, L.S.; Zhang, P.; Hao, H.; Zhu, L.; Toomey, N.L.; Lee, M.Y.W.T.
Nucleic Acids Res. 20, 735-745, 1992
A. Title: Molecular cloning of the cDNA for the catalytic subunit of human DNA polymerase
A. Reference number: S35455; PMID:92178967; PMID:1542570
A. Accession: S35455
A. Molecule type: mRNA
A. Residues: 1-29, 'R', 31-118, 'H', 120-172, 'N', 174-471, 'Y', 473-775, 'G', 777-1107 <YAN>
A. Cross-references: EMBL:M81735; NID:J9181621; PIDN:AAA53768.1; PID:J9181622
C. Gene: POL3
A. Gene: GDB:POLD1; POLD
A. Cross-references: GDB:129089; OMIM:174761
A. Map position: 19q13.3-19q13.3
C. Superfamily: herpesvirus DNA-directed DNA polymerase
C. Keywords: DNA binding; DNA replication; nucleotidyltransferase; nucleus; zinc finger

Query Match 18.2%; Score 760.5; DB 1; Length 1107;
 Best Local Similarity 28.4%; Pred. No. 3.2e-38;
 Matches 252; Conservative 145; Mismatches 324; Indels 165; Gaps 31;

17 GKEPQVITIGIAENGERVYLIDRSFRYPYFALLARQADPKVOAQRIR---ALSRRK--- 69
 120 GSVPLVIRAGVTDSEFSCCHIHGFAYFTTPAPGPFGEHMGDLQELMLAISRSRGG 179
 70 -----SPITGVDDKRX---YFGR-PRRYLRITVLRPEAVRELVKQVGVDEVDV--- 116
 180 RELTGPAVLAVALCSRESMEFGHGHGSPFLRIYALPRLVAPARRLLEGGIRVAGIGTP 239
 117 -----LEADIRPMRYLIDHDLFPFTYRYAELEPNKGFVNDKYVL-----YKSRP 164
 240 SPAPYANVDFEIRRMVDTDIVGCNMLEPAGVYALRLKERATQCOLEADVLSVDVSHP 299
 165 EPIYGEALAPTKLPDLRIILAFDIEVYSKQG-SPRPRDPYIV---AVKTDGDEVLEFLA 220
 300 PEGPWGRINP-----LRTLSFDIECGRGRGIPPEPRDPVQIQSLSLRNGDEPEPLRLA 354
 221 -----RGQDKRPIREFRYVYKAYDPDIIYGVNNHHPWVYLLRRA--- 261
 355 LTRPAPPLAKVNSYKEDL--IQANSTRIRMDPEVLTGRIQVDFLPLISRAQT 412
 262 -----RLIGKLDV-----TRRVGAEPPTSVHGHVSVGRINVDLVYAEEMPE 305
 413 LKVQTFPLGRVAGLCSNIRDSFSQSKQGRDRDTRV---VSMVGRVQMIMDLQVLLREYK 468

QY 306 IKISLEVA-BYIGVMKSERVIINMWEIPDYMDPKRPLLQYARDVR---ATYGL 361
Db 469 ILSHTLNANVSHFIFGEQKEDVQHSI-----ID-----LQNGNDQTRRLAYCYL 513
QY 362 AEKILPF-----AIOISYVTGLPLDQVAMS VGFLREMYLIRAAFOMKELVNR 410
Db 514 KDAVLPILRLERLWLVNAVEMARVGVPLSYLSRGOQVAVSOLRQAMHEGLMPVV 573
QY 411 VERPEETRGALVEPLAGVHE-NIAVLDPSSMYNIMIKNGBDITVRPG--EKCGEC 467
Db 574 KSEGEDYTGATVLEPLKGYDVPIATLDFSSLYPSIMAHNLCTYTLIRPGTAQKLG-- 631
QY 468 GCWEAPEVKHFRRCPPG--FFKT-----VLERLLELRKRVRAEMKKYPPSPPEYR 516
Db 632 -----LTEDQFIRTPGTDEFVKTSVRKGLLPQILENLSARKAKAKELAK-ETDPLRQ 684
QY 517 LLDERQALKVLANASVGYMGSGARMYCRECAVATAMGRHLI-----RTAINIA 567
Db 685 VLDERQALAKVSANSVYGTGAQVGLPCLEPSQSVTGFGRQMI EKTQVLESKYTVENG 744
QY 568 RKGLKVIYDPTDSLFTYDPEKVENFIKIKEELG-----EIKLEKYTKRLFFT 618
Db 745 YSTSAKVYVGTDSVMCRFGVSSVAEAMALGREADWVSGHFPSPIRLFEFEKYFP-YLL 803
QY 619 EAKRVAAGIL-----EDGRIDIVGFEAVRGDMCELAKEVOTKVVEIVLKTSEVNAKEY 672
Db 804 ISKRKYAGLFFSSRPDAHDRWDCKGLEAVRDNCPVANLVTAISIRLLIDRDPEGAVAH 863
QY 673 VRKIVKELEBGKVPLEKLVIMKTLISKLEBYTTEAPHVAAKRM--LSAGYRVSPGDKIG 730
Db 864 AODVISPDLCHRIDISQLVITKEILTBAASDYAGQAHTELERMRKRDPGSAPSLGDRVP 923
QY 731 YVIV---KGGRIQORAPYFMWKPDSQIDVTYVVDHQIIPALRI 773
Db 924 YVISAAGVAAVYKMSBDPLFVLEHSLPIDTQYYLEOQLAKPLRI 969

Search completed: Apr 11 23, 2003, 07:29:40
Job time : 61 secs

Delaval, Jan

92138

From: Haddad, Maher
Sent: Wednesday, April 23, 2003 7:17 AM
To: Delaval, Jan
Subject: 10/034,849

Mahe Haddad, 1644
9E12/9D08
306-3472

Jan, .

Please search
1) SEQ ID NO:2 and
2) oligo of seq id no. 2.

Thanks-Mahe

Jan Delaval
Reference Librarian
Biotechnology & Chemical Library
CM1 1E07 - 703-308-4498
jan.delaval@uspto.gov

5

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 07:28:09 ; Search time 15 Seconds
(without alignments)
2220.366 Million cell updates/sec

Title: US-10-034-849-2

Perfect score: 4188
Sequence: 1 MTEWFTVLDSYEYVGEK.....LKASATGQKTLFPLAKSK 803

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2434	58.1	784	1	DPO2_AERPE
2	1581	37.8	781	1	DPO1_ARCEU
3	1303.5	31.1	763	1	DPO3_SULSH
4	1287	30.7	764	1	DPO3_SULSO
5	1245.5	29.7	773	1	DPO1_THESQ
6	1236.5	29.5	775	1	DPO1_PYRFU
7	1207.5	28.8	771	1	DPO1_PYRAB
8	1197.5	28.6	775	1	DPO1_THES9
9	986	23.5	824	1	DPO1_METVO
10	965.5	23.1	1312	1	DPO1_PYRSD
11	955.5	22.8	1235	1	DPO1_PYRHO
12	915	21.8	586	1	DPO1_METH
13	845	20.2	1523	1	DPO1_THERM
14	776	18.5	1671	1	DPO1_PYRKO
15	771.5	18.4	1086	1	DPO1_SCHPO
16	765	18.3	1105	1	DPO1_ORYSA
17	760.5	18.2	1107	1	DPO1_HUMAN
18	757.5	17.9	1106	1	DPO1_BOVIN
19	749.5	17.9	1699	1	DPO1_THES8
20	749	17.9	1103	1	DPO1_CANAL
21	746.5	17.8	1103	1	DPO1_MESAU
22	743.5	17.8	1105	1	DPO1_MOUSE
23	741.5	17.7	1088	1	DPO1_SOYNS
24	737.5	17.6	1081	1	DPO1_ARATH
25	736.5	17.6	1103	1	DPO1_RAT
26	713	17.0	11702	1	DPO1_THELI
27	705	16.8	1094	1	DPO1_PLAFK
28	691.5	16.5	1097	1	DPO1_YEAST
29	670	16.0	1668	1	DPO1_METH
30	670	16.0	1634	1	DPO1_METH
31	658.5	15.7	1829	1	DPO1_THES1
32	655	15.6	1092	1	DPO1_DROME
33	647.5	15.5	872	1	DPO1_SULOH

34	628.5	15.0	875	1	DPO1_SULAC
35	625.5	14.9	959	1	DPO1_AERPE
36	598	14.3	1081	1	DPO1_CABEL
37	594.5	14.2	882	1	DPO1_SULSO
38	579.5	13.8	1462	1	DPO1_HUMAN
39	578	13.8	1465	1	DPO1_MOUSE
40	568	13.6	1451	1	DPO1_RAT
41	539	12.9	3130	1	DPO1_HUMAN
42	526	12.6	3122	1	DPO1_MOUSE
43	524.5	12.5	1015	1	DPO1_EBV
44	522.5	12.5	1339	1	DPO1_TRYBB
45	515	12.3	1009	1	DPO1_HSVSA

ALIGNMENTS

RESULT 1
DPO2_AERPE STANDARD; PRT; 784 AA.
ID DPO2_AERPE
AC 093746;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DNA polymerase II (EC 2.7.7.7).
GN POLB OR APE2098.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcaceae;
OC Desulfurococcaceae; Aeropyrum.
% NCBI_TaxID=56636;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RA Ishino Y., Cann I.K.;
RT "Isolation of the genes encoding two alpha-like DNA polymerases from Aeropyrum pernix.";
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RA MEDLINE=9310339; PubMed=10382966;
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Hatake Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Anai A., Kosugi H.,
RA Hasegawa A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamaya M., Masuda S., Funahashi T., Tanaka T., Kudo Y.,
RA Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
-1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate + {DNA}(n).
-1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL, AB017501; BAA75663.1; -;
CC EMBL, AP000063; BAA81109.1; -;
CC HSSP, P56689; ITGO.
DR InterPro: IPR002064; DNA_pol_B.
DR InterPro: IPR004578; Pol2.
DR Pfam, PF00136; DNA_pol_B; 1.
DR Pfam, PF03104; DNA_pol_B_exo; 1.
DR PRINTS; PR00106; DNAPOLB.
DR SMART; SM00486; POLB; 1.
DR TIGRFAMs; TIGR00592; pol2; 1.
DR PROSITE; PS00116; DNA_POLYMERASE_B; 1.

KM Transferrase; DNA-directed DNA polymerase; DNA replication;
 KM DNA-binding; Complete proteome.
 FT CONFLICT 104 104 E -> K (IN REF. 1).
 FT CONFLICT 306 319 KIGERYLEMOIG -> R (IN REF. 1).
 FT CONFLICT 355 355 E -> Q (IN REF. 1).
 FT CONFLICT 754 759 LRLOY -> SHTSSN (IN REF. 1).
 FT CONFLICT 763 763 T -> I (IN REF. 1).
 SQ SEQUENCE 784 AA; 89639 MW; F36311C4063B2075 CRC64;

Query Match: 58.1%; Score 2434; DB 1; Length 784;
 Best Local Similarity: 59.1%; Pred. No. 3, 7e-146;
 Matches 469; Conservative 113; Mismatches 189; Indels 22; Gaps 5;

Qy 17 GKEPOVITWICINENGERVVLIDRSFRFYFALTAAGADKQVACIRALSRPSPIIGV 76
 Db 3 GSTFVITLMGRADGSRVAVYGFRRFYFALDGSVGLDOLAAIRLSFSSPLSVE 62
 Qy 77 DDKKTFGRPRVLRIRITVLEAVREYRELKVNQGVEDVLEADIRFAMRYLIDHDLFPF 136
 Db 63 RVRRRRIGREVALKVTTLVPASVREYREVRRLGCVBVDLEADIPFALPFIIDFNLPM 122
 Qy 137 TWYVBAEPLBNKMGFRVDKYYLV-----KSRPELYGEALAPTKLPDIRIAPDI 187
 Db 123 RMVVAEVRVAAPHGYSVDRAVLTSGDIRREDTRIODEPLKG-----LRVMAFDI 172
 Qy 188 EYVSKQGSRRPRPDVVIYIAVKTDDGDEVLFIAEGKDDKPIREFVEYVKYDPDIIVG 247
 Db 173 EYVSKMRTDPKDDPVIMIGIQAGSRIELEBEDSDCKVIAGFVRYSIDPDVIVG 232
 Qy 248 NNHFPMPYLRRARILIGIKDYTRRGAPETTSVHGHSVSGRLANDVYAEEMPEIK 307
 Db 233 NQRFPMFPLVBARAVLCKLAVRR-SVEPQGLIGHYSVSGRLANDLDFABELHKA 291
 Qy 308 IKSLAEVAYLGWKKSERVYINMWEIPDWDPKRPLLOYADRYATYGLAEKIIP 367
 Db 292 VKLEAVADYLDGVKIGRVTLREWMQIGETWDDPSKRETLKRLDQVRSWTGLAEKLP 351
 Qy 368 FAIQLSYVTGLPDDVGMASVGFLEWYILRAAFKKKELVPRVREPEYRGAIVLEP 427
 Db 352 FGALISQVSGPLDDVMAASVGFLEWRLIREAKKLGELVPRVRSRGAQIVLRPK 411
 Qy 428 RGVHENIADVSSMYPNIMKYNVGPDTLVPRGECGCGCMEAEVKIRFRCPGPF 487
 Db 412 PGVHEDIADVDSMYPNIMKYNVGPDTLVPRGEEYGEVEYTAPEVGHKFRKSPGPF 471
 Qy 488 KTVLERILELRKRVBAEMKKYPPDSPEYRLDERQALKVLANASVGMGSGARMYCRE 547
 Db 472 KTLERFLSMRQIRSEMKKHPDPSPEYRLDERQALKVLANASVGMGSPHARYCRE 531
 Qy 548 CAAVATVAMGRHLIRTAIINIAKGLKVIYGDTSLFVYTDPEKVENFKIIXELGPEIK 607
 Db 532 CAAVATVAMGRSIIIRTAIKKAGELDELVIYDTSLFVKNDEPKVERLIRVEBEELGFDIK 591
 Qy 608 LEVYVRLFTFAKKRYAGLLEDRIDIVGEFANRGDMCELAKEVQTVVEYVLAKSEVN 667
 Db 592 VDVYVRFVETFAKKRYAGLTVGKIDVGEFANRGDMSELAKETQFVAVELVAKSGVD 651
 Qy 668 KAVEYVRIKVELEEGKPIELKVIYKTSKLEETTEAPHVAARMLSAGRVSPD 727
 Db 652 EADVYVRNIIETLRGQVDMKLVIMKTLTRPPSMYEAQPHVTAALLMRBAGIKYEPBA 711
 Qy 728 KIGYVIVKGGRIISQRAWPYFVVDPSQIDVYVYDHOIIPALALIIIGEGITEKLLAS 787
 Db 712 KIGYVIVKGGSGPLVTRAKPYEMA-SKEEVVVEYVDKQVAPALRIIQYGVTEKLLKGG 770
 Qy 788 ATGQKTLFDFLAK 800
 Db 771 GR-QSTLDFMR 782

RESULT 2
 DPOL_ARCFU
 ID DPOL_ARCFU STANDARD; PRT; 781 AA.

AC O29753;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE DNA polymerase (EC 2.7.7.7).
 GN POL OR POLB OR APO457.
 OS Archaeoglobus fulgidus.
 OC Archaeae; Euryarchaeota; Archaeoglobi; Archaeoglobales;
 OC Archaeoglobaceae; Archaeoglobus.
 OX NCBI_TaxId=2237;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
 RX MEDLINE=89049343; PubMed=9389475;
 RA Kleink H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
 RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
 RA Richardson D.L., Kervavage A.R., Graham D.E., Kyriades N.C.,
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
 RA Kitzness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
 RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodok A., Zhou L.,
 RA Cotton M.D., Spriggs T., Artach P., Kaine B.P., Sykes S.M.,
 RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
 RA Venter J.C.;
 RA "The complete genome sequence of the hyperthermophilic, sulphate-
 RT reducing archaeon Archaeoglobus fulgidus."
 RL Nature 360:384-370(1997).
 CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
 CC + (DNA) (N).
 CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AE001070; AAB90741.1; -.
 CC
 CC DR HSSP; P56689; ITGO.
 CC DR TIGR; AF0497; -.
 CC DR InterPro; IPR002064; DNA_pol_B.
 CC DR InterPro; IPR004578; Pol2.
 CC DR Pfam; PF00136; DNA_pol_B_1.
 CC DR Pfam; PF03104; DNA_pol_B_exo_1.
 CC DR PRINTS; PR00106; DNAPOLB.
 CC DR SMART; SM00485; POLBc; 1.
 CC DR TIGRFAM; TIGR00592; pol2_1.
 CC DR PROSITE; PS00116; DNA_POLYMERASE_B_1.
 CC KM Transferrase; DNA-directed DNA polymerase; DNA replication;
 KM DNA-binding; Complete proteome.
 SQ SEQUENCE 781 AA; 89850 MW; 761C7B18FCC61B8B CRC64;

Query Match: 37.8%; Score 1581; DB 1; Length 781;
 Best Local Similarity: 42.5%; Pred. No. 1, 3e-93;
 Matches 350; Conservative 133; Mismatches 273; Indels 68; Gaps 18;

Qy 1 MTEVFTVLDSSEYVNGKEPOVITWICINENGERVVLIDRSFRFYFALTAAGADKQVAC 60
 Db 1 MEREGMLIDADYETIGGKAVVRLMCKDDG-IFVAYDYNDFPFYVY---GVDEILKN 56
 Qy 61 RIRALSRPSPIIGVEDDKRYGFRPRVLRIRITVLEAVREYRELKVNQGVEDVLEADIRF 120
 Db 57 AATSTRREVYIKLSFEKQKLTIGREVEGYIVAAHHPGVRLKADYLSOG---DVREAD 113
 Qy 121 IRPAMEYILDHDLPEFTYRVRVAEPLBNKMG-----FRVDKYYLKSRRPELYGEALAP 175
 Db 114 IPRYRILIDKDLACMGDGIATIGE---KGGGVYSYKIEKV-----ESTIPM 157
 Qy 176 KLPDLRIAPDIEVYVSKQGSRRPRPDVVIYIAVKTDDGDEVLFIAEGKDDKPIREFVEY 235

Db 158 EPELKMVLDPDCEMLSSFGMEPEKDPPIVIVSKTNDDEIILTG---DERKIISDFVKL 214
 Qy 236 VKRVDPIIVGYNNHPPWPLRLRRARILGKLDVTRVGAPEPTSYH--GHVSVPRGLN 293
 Db 215 IKSDPDIIVGNDAPFMPPLRKAEKRWNIPLD---VGRDGSNVVFRGRPKITGLN 270
 Qy 294 VDLVYAEEMBEIKIKSLAEVAYLVGVAKSERVILNMWEIPDYWDDEPKKPLLLQVARD 353
 Db 271 VDLVIMRISDIKIKLENAEFLGT--KLEIADIKAQDLYRWSRREKCK-VLNTARQ 327
 Qy 354 DVARTYGAEKILPPAQLSVTGLPLDQVAGSVGFLEVYLIRAPKMKELVPRVER 413
 Db 328 DAITYTLAKELIPMHVELSKIRLPVDDVTRMGKQVMDLLSEAKKIGELIAPNPEH 387
 Qy 414 PEITYRGAIYIEPLRGVHENAVLDFSSMYPNIMIKYVNGVDTLVREPKGEGCCWAP 473
 Db 388 -ASBYEGAFVLEPERGHEENACLDPSMYPNIMAFNISPDY----GGRD-DCEAP 440
 Qy 474 EVKRFRCRCPGFETVLERLELRKVRAMKKYPPDSPERYLLDEROKALKYLANASY 533
 Db 441 EVGHKFKSPGPFKRIIRMLIEKRELKVELKSLSPSSSKLDDIKQTLKLTNSFY 500
 Qy 534 GYMGSGARWYCRCAKAVTAMGRLIRTAIRAKGLKVIYGDTSLPVT--YDPEK 590
 Db 501 GYMGMLARWYCHPCAETAMGRHIFRTSAKIAESMGFKVLYGDTDSIFVTKAGMTKD 560
 Qy 591 VENIKIKELGFEIKLEKYKRLFTFAKRVAGLLEDRIDIVGEAVRGWCELAK 650
 Db 561 VDRILDKHEELPIQIEVDEYYSALPFEV-KRYVAGLTEDRLVVGLEVRGRWCELAK 619
 Qy 651 EVQTKVEIVLKTSEVNAKAVEYRKIVALEEGKVPLEKLYWKTLSKRLSEYTTAPHV 710
 Db 620 KVGQEVLEVILKEKNPEALSLVQVILRIKEGKVSLEEVYIKGLTKPKESYMAHV 679
 Qy 711 VAAQMLSAGRYVBPQDKIYVIYKGGRIQORAMPYFVKD-----PSQ 755
 Db 680 KALKAKEMGIIVPVSSKIGVIYVSGNIGDRAPVILIDIDFQENLRITKSGIETIK 739
 Qy 756 IDVYVYVDOHIIIPALRIIGYFITEKKLKASATGQKTLFPFLA 799
 Db 740 LDKQYIDNOIIPSVLIRLFRFGYTEASLKSS--QMSLDSFFS 781
 RESULT 3
 DPO3_SULSH STANDARD; PRT; 763 AA.
 AC 005706;
 ID 15-DEC-1998 (Rel. 37, Created)
 AC 15-DEC-1998 (Rel. 37, Last sequence update)
 DE 15-JUN-2002 (Rel. 41, Last annotation update)
 GN DNA polymerase III (EC 2.7.7.7) (DNA polymerase B3).
 OS Sulfolobus shibatae.
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 OC Sulfolobus.
 OX NCBI_TaxID=2286;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97252493; PubMed=9098062;
 RA Edgell D.R., Klenk H.-P., Doellittle W.F.;
 RT "gene duplications in evolution of archaeal family B DNA
 RT polymerases";
 RT J. Bacteriol. 179:2632-2640(1997).
 CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
 CC + (DNA) (N).
 CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>)

CC or send an email to license@isb.ch.
 CC -----
 Db EMBL: U92874; AAB53089.1; -
 Dr Interp: IPR02064; DNA_pol_B.
 Dr Pfam: PF00136; DNA_pol_B_2.
 Dr SMART: SM00486; POLBc_1.
 Dr PROSITE: PS00116; DNA_POLYMERASE_B; FALSE_NEG.
 Km Transferrase; DNA-directed DNA polymerase; DNA replication;
 Km DNA-binding; 763 AA; 88196 MW; 7A7EFD9A5B9305E2 CRC64;
 SQ SEQUENCE
 Query Match 31.1%; Score 1303.5; DB 1; Length 763;
 Best Local Similarity 35.2%; Pred. No. 7e-76;
 Matches 282; Conservative 164; Mismatches 309; Indels 47; Gaps 14;
 Qy 6 FTVLDSSEYVAGKEPOVITWIGLNGERVLLIDRSFRPYFALLAPGADQVQVORIAL 65
 Db 5 FFLIDFSEYIKDNIPLIYIWSIDEGNSCVVERNFPEYVYVEGND--EILENI--- 59
 Qy 66 SRKSPILIGVEDDKRYKFGPRVLRIRITVLPEAVREYRELKVNQGVDEVLADIRFAM 125
 Db 60 -RKNCEVLLITTKVKRYLGNVDALLVQFTPTQIRCREKISRINGIKSIFPADIRFTM 118
 Qy 126 RYLIDHLPFTWYVYAEAELENMGFRVDKYLIVKSRPEPLVGEALAPYKLPDLRIAP 185
 Db 119 RYSIDPDLRFTFKAEVSEVKE-GFRACKVILDKILSHYG-----KIFELRAIGI 171
 Qy 186 DIEVYSKQSPREPERDVIVIAVKTDDGDEVLPIAGKDRKPIREFEVYKRPDDIIV 245
 Db 172 DFOIYSKYGSLNPRKDIIVLTLMSKSGSWQFSLDSMDLKIIRKVDYILNYDDIIV 231
 Qy 246 GYNNNHDMYLLRRARILGKLDVTRVGAPEPTSYHGVSPGRANDLVYAAEMPE 305
 Db 232 VFDVDFHMYITITERANSLGKVIDIGKIGSEVSGTYGHYISIGRANDLVGLMNERL 291
 Qy 306 IKIKSLAEVAYLVGVAKSERVILNMWEIPDYWDDEPKKPLLLQVARDVARTYGLAEKI 365
 Db 292 TGHIDILEVANYIGISK--RDSLWYETISRYNDDENRDLVQYSLNKAISYLLGNFL 349
 Qy 366 LPPAQLSVTGLPLDQVAGSVGFLEWYLIRAPKMKELVPRVERPEYTR-GAIVL 424
 Db 350 LSPYSELVKIIGLPLDGLSAVSGNRIEASLIRTAASEELIPIRMDNPRSSKIKKTVI 409
 Qy 425 EPLRGVHENAVLDFSSMYPNIMIKYVNGVDTLVREPKGEGC-GCEAPAEVKAHRRCP 483
 Db 410 EPKIGIYSDVYVLDISSVYLSVIRKFNISPDILV-----GQCDCYVSTISNYKFKEP 464
 Qy 484 PGFEKTVLERLELRKVRAMKKYPPDSPERYLLDEROKALKYLANASGYMGSGAR 543
 Db 465 SGLYKTFLELSNIQD-----TRKSVIIEELSSFYDYIHWINSRW 505
 Qy 544 YRECAKAVTAMGRHLIRTAIRAKGLKVIYGDTSLPVT-YDEKVENFKIIEEL 602
 Db 506 YSRHIAAVDELSEYELGKLVITDILKNSGFEVILLANPLVYVKGSGDKNELLFKINSLV 565
 Qy 603 GFBKLEKVTYKRLPFTFAKRYAGLLEDGRIDIVGEAVAGDMCEIAKEVQTVYVEIVLK 662
 Db 566 DLMKAKKIRYSLIIL-GNDRYAGLLEGDKIDIRIGKEDRDICEIVRNKRYAVEIILI 624
 Qy 663 TSEYKAVEYVRKIYVLEEGKVPLEKLYWKTLSKRLSEYTTAPHVYAKMLASGR 722
 Db 625 SKDYKAVKAVKANVILKRGGEFDIGELITWVHIEKDFSYDKQLPVPVAAARAIOSGYL 684
 Qy 723 VSPGDKIGVIYKGGRIQORAMPYFVKDPSQIDVYVYVDOHIIIPALRIIGYFITEK 782
 Db 685 ISKDSRIGYLVIGHGSHVDRAPFPFVKEKNRIDIEYVD-QLNEBSLKVLPPLGVSEE 743
 Qy 783 KLRASATGQTLTD-PLAKSK 803
 Db 744 SLK--KTNIYDILDMFGASKK 763
 RESULT 4

DP03_SULSO
 ID DP03_SULSO STANDARD; RT; 764 AA.
 AC P95979;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE DNA polymerase III (EC 2.7.7.7) (DNA polymerase B3).
 GN DP03 OR SS00081 OR C04041 OR C05_029.
 OS Sulfolobus solfataricus.
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 CC Sulfolobus.
 CX NCBI_TaxId=2287;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 35092 / DSM 1617 / P2;
 RX MEDLINE=9705332; PubMed=889719;
 RA Senen C.W., Klenk H.-P., Singh R.K., Allard G., Chan C.C.-Y.,
 Liu Q.Y., Penny S.L., Young F., Schenk M.E., Gaasterland T.,
 Doolittle W.F., Ragan M.A., Charlebois R.L.;
 "Organizational characteristics and information content of an
 archaeal genome: 156 kb of sequence from *Sulfolobus solfataricus*
 P2";
 RL Mol. Microbiol. 22:175-191(1996).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 35092 / DSM 1617 / P2;
 RX MEDLINE=21332296; PubMed=1147726;
 RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
 Aways M.J., Chan-Weber C.C.-Y., Clausen I.G., Curtis B.A.,
 De Moers A., Eruso G., Fletcher C., Gordon P.M.K.,
 Heilmann de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
 Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
 Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
 Garrett R.A., Ragan M.A., Senen C.W., Van der Oost J.;
 "The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
 RN (3)
 RP DISCUSSION OF SEQUENCE:
 RX MEDLINE=97252493; PubMed=9098062;
 RA Edgell D.R., Klenk H.-P., Doolittle W.F.;
 "Gene duplications in evolution of archaeal family B DNA
 polymerases";
 RL J. Bacteriol. 179:2632-2640(1997).
 CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
 + (DNA)(N).
 CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: Y08257; CA69570.1; --
 DR EMBL: AB006646; AA04041.1; --
 DR InterPro: IPR002064; DNA_POL_B.
 DR Pfam: PF00136; DNA_POL_B.1.
 DR SMART: SM00486; POLBc.1.
 DR PROSITE: PS00116; DNA_POLYMERASE_B; FALSE NEG.
 DR Transferrase; DNA-directed DNA polymerase; DNA replication;
 KW DNA-binding; Complete proteome.
 SQ SEQUENCE 764 AA; 88378 MW; 1AA325C7D9A4D672 CRC64;

Query Match 30.7%; Score 1287; DB 1; Length 764;
 Best local similarity 34.8%; Pred. No. 7,9e-75;
 Matches 280; Conservative 167; Mismatches 305; Indels 52; Gaps 16;

OY 6 FTLTSSVYVGKKEPOVIVICIAENGERVYLIDRSRRPYFALLAPADKQVAGQIRL 65
 DB 5 FTLDSTYIKONTPLVIVINSVDDEGNSVVDNRRFPYIYIE--GNEEITENI--- 59

OY 66 SRPKSPIIGVEDDKRYFGRRPRVLRIRTVLBEAVRELVKAVDGVEDVLBADIRFAM 125
 DB 60 -KNCCEALQITKVKKYLGNIVDALLIOTSTPQIKKCEKSEIINNKKIPADIRYTM 118
 OY 126 RYLIIDDLPPFWYVEAEPLENKMGPRVDKQVYLVSRRPEPLYGALAPKRLDLILAF 185
 DB 119 RYLDLDLAPFWYFAEYNEVKPD-GRRTKAVIYDKLISHYEG-----NMELETTIV 171
 OY 186 DLEVYKSGSPREEDPVIYAVTDDGDEVFLIEAGDDPKRPREVEYKXPDDIIV 245
 DB 172 DFOIYKSGSLNPKDFIVMSLKSQPMQSLDEGIDDLKIRRVYDIILYVDDIIF 231
 OY 246 GYNNHFDPWYLLRRARILGKLDVTRVGAEPPTTSVGHVSVFGRILVLDLYAEEMPE 305
 DB 232 VYSDLLPWRKYITERPSSLVGKIDIGKIGSEVSGTGVSGIRLNVDTGLVNER 291
 OY 306 IKISSLAEVAYELGVKKSERVLIINWELIPDWDPKRPPLLQYARDVVRATYGLAEKI 365
 DB 292 LGHVLDIVSNYLG--SPSRYSFKWYELSRWDEKRRIRIRESIENRSIYLLGNVYL 349
 OY 366 LPEAIOLSYVTLPLDQVGAHSVGRLEWYLIRAAFKMKELVPRVPRPE--ETYGALIV 423
 DB 350 LSTYSELVYVGLPLDKLSVASMGNRIETSLIRTAIKSGELIPRMONPRPKIKNII 409
 OY 424 LEPRLGVHENIIVLDFSGMYPMIKYNGPPTIVRPGKGGCEGCEAEVGHRRRCPC 483
 DB 410 IQPRVGIYTDVYVLDISSVSLVIRKFNAPDLVYK--EQCDD--CYSSPISNYKKKEP 465
 OY 484 PGFKTVLERLLLELKRVRAMKYPDPDEPRILDEROKALVLANASGYNGSGARW 543
 DB 466 SGLKTFLELSNV-----DSKKIKVIEE-----LISSPDYHVAWNAW 506
 OY 544 YCECAKAVTAGRHLITAINARKLGVYIGDTSLPYT-YDPERVENFIKIIEEL 602
 DB 507 YSEIISAPADEFSNEIIFIIIDLKSSGLDVIANDLLIFVYGSRKRVNELLTKINSLY 566
 OY 603 GFEIKLEKYKRLFTFAKKRYVAGLLEDGRIDIVGFEAVRGD---WCELAKEVQTKVE 658
 DB 567 NLVVKVIRYKSLVLVD--NNRYVAGLSEGDKIDI---ARKSEEDMNICEIARINIKKIE 621
 OY 659 IVLKTEVNAKVAEYVRKIVKELGKVPLEKIVMTLSKLEETEAHHVAARML 718
 DB 622 EILISDVKAIKLVKSVIVLRGEPDNEELITAKIERDLENNQQLPFTVARKAIQ 681
 OY 719 AGRVSPGKIGIVYVKGCGRISGRAMPYMWDSQIDVYVYDHOIIPAAIRIIGYRG 778
 DB 682 SGLIKSDKIKIYVVKGLGFLNDRAPFTYKSKRIDIEYVD-QIFRETIKLAKPG 740
 OY 779 ITEKLI-KASATGCTLPFLAKK 801
 DB 741 VNESLAKTNITDIIDLFGASKKK 764

RESULT 5
 DBL THEGO STANDARD; RT; 773 AA.
 ID DBL THEGO
 AC P56689;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE DNA polymerase (EC 2.7.7.7) (TO POL).
 GN POL OR POLA.
 OS Thermococcus gorgonarius.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 CC Thermococcus.
 CX NCBI_TaxId=71997;
 RN (1)
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
 RX MEDLINE=99192304; PubMed=10097083;
 RA Hopfinger K.-P., Eichinger A., Engn R.A., Laue F., Ankenbauer W.,
 Huber R., Angerer B.;
 "Crystal structure of a thermostable type B DNA polymerase from
 Thermococcus gorgonarius";

DR PROSITE; PS00116; DNA POLYMERASE B; 1.
 KM Transferrase; DNA-directed DNA polymerase; DNA replication;
 KM DNA-binding; Complete proteome.
 FT VARIANT 532 532 R -> S (IN STRAIN GE23).
 FT VARIANT 553 554 HE -> PN (IN STRAIN GE23).
 FT VARIANT 756 756 K -> R (IN STRAINS GE23 AND ST 855).
 FT CONFLICT 263 263 V -> T (IN REF. 1).
 FT CONFLICT 277 277 A -> T (IN REF. 1).
 FT CONFLICT 281 281 A -> V (IN REF. 1).
 FT CONFLICT 320 320 P -> S (IN REF. 1).
 FT CONFLICT 339 339 Q -> H (IN REF. 1).
 FT CONFLICT 359 359 R -> T (IN REF. 1).
 FT CONFLICT 391 391 K -> N (IN REF. 1).
 SQ SEQUENCE 771 AA; 89496 MM; 110A67045A8A5522 CRC64;
 Query Match 28.8%; Score 1207.5; DB 1; Length 771;
 Best Local Similarity 35.7%; Pred. No. 9.7e-70;
 Matches 292; Conservative 150; Mismatches 285; Indels 91; Gaps 21;
 8 VLDSSEYVVGKEPOVYIIGIAENGERVVLIDRSFRPYFALLAPGADPKOY----AQRI 63
 2 IILDYITDEGKPIIRIK-KEGEFKVEYDRTFRPYIYALKDSDAIDEVKKITAEHSG 60
 64 ALSPPKSPDIIGVEDDKRYGPRPRVLRIRTVLPEAVREYRELKVNVDGVEDLEADIR 123
 61 KIVR-----ITEVEKVKQKFLGRPIEWMKLYLEHPQDVPAIREKHREHNAVVDIEYDIP 116
 124 AMRVLIDHDLFPFTWYRVEAPLEENKGFVVDKYLVKSRPEPLYGELAPTKLPDLRIL 183
 117 AKRVLIDKGL-----TPMEN-----EELTF 138
 184 AFDIEVYVSKGSPRPERDPVIVAVKTDGDEVL-----FLAEGKDDKRPREFEY 235
 139 AVDIETLYHEGE-EGGKPIIMISYADEGAKVITWKSIDLPYEAVSSEREMIKRLVK 197
 236 VKRYPDIIIVGNNHFPYLLRRARILGKLDVTRVGAEPITSVGH---VSVPGL 292
 198 IREKDPDVIYVNGDNFPFYLKRAEGLKPLGRD-NSEPKQMGDGLAIVEIKRI 256
 293 NVLDYVAEENPEIKISLEVAEVLGYMKKSERVIINMWEIPDYDDPKRPLLOVAR 352
 257 HFDFPVLIRRTINPTLYLEAVYEAI-EGSKSEKVVAAH-EIAEAMWEGKGLERVAKYSM 313
 353 DDVATYGLAKILPFAITQSVTGLPLDOYGAMSVGRLEEMYLIRAFKKKELVPRNVE 412
 314 EDKATVPELGEPFPMELKRLVGPWADSRSSSTGLVEMFLRKAKYERNEELAPNRPD 373
 413 RPE-----ETTRGAILVLEPRGHEINIAVLDFSSMYPNIMIKYVGPDTLVRPEKCGE 466
 374 EREYERRLRESEYEGVYKEPEKGLMEGLVSLDFNSLVSIIITTNVSPDTLNR--ENCKE 431
 467 CGCEAPVVKRRFRRCPPGFKTVLERLLELRKRVRAEMKYPDPSPRYLLDEROKALK 526
 432 YDV--APQVGRFCDFPGFIPSLGILNLEEROKIKKMKKE-SKQPVKSKLLDYRQRAIK 488
 527 VLAAVSYGVMGSGARWYCRCAKAVTAMGRHLIRTAIINARKGLKTIYDQTSLEFTY 586
 489 ILAASYGYTGAKARMTCKECAESVYAMGQYIDLVRRELESREFKVLIDTQGLVATI 548
 587 D-----PEKVENFIKIKEELG--FEIKLEKVKRLFTFAKKRYAGLLEDGIDIVG 637
 549 PGAGHEBEIKAKLKVEYINSLKPLGLELEVBGFARGFV-TKKKVALIDEBEGKIVRG 607
 638 PEAVRGDMCELAKYQTVNVEIVLKTSEVNTAVVEYRKIVLEBEGKPIEKLVIWKTLS 697
 608 LEIYARDMSKIAKETOAKVLAAILKNGVNDVAVKIVKEVTEKLSKYEIIPPEKLVYEQIT 667
 698 KRLSEYTEAPHVVAARAKMLSAGYVSPGDKIGYIVVGGGRISGRAMPYMWK--DSS- 754
 668 RPLSEYKAIIGHVAVAKLAKGVAKVKGPMYIGIIVLEGGDGPISGDA--IAIEFPDKK 724
 755 -QIDVTYVVDHQQIIPALRLIIGYFGITEKUKASATGQ 791

DB 725 HKYDAEYIENQVLPAYERILRAFGRKEDLKYYOKTKQ 762
 RESULT 8
 ID DPOL_THES9 STANDARD; PRT; 775 AA.
 AC 056366;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE DNA polymerase (EC 2.7.7.7).
 GN POL OR POLA.
 OS Thermococcus sp. (strain 90N-7).
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Thermococcus.
 OC NCBI_taxid=103799;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96224274; Pubmed=8643567;
 RA Southworth M.W., Kong H., Kucera R.B., Ware J., Jannasch H.W.,
 RA Perler F.B.,
 RT "Cloning of thermostable DNA polymerases from hyperthermophilic marine
 RT Archaea with emphasis on Thermococcus sp. 90N-7 and mutations
 RT affecting 3'-5' exonuclease activity."
 RL Proc. Natl. Acad. Sci. U.S.A. 93:5281-5285 (1996).
 CC -I- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
 CC + (DNA) (N).
 CC -I- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.
 CC CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC CC
 DR EMBL; U47108; AAA8769.1; -.
 DR HSP; P56689; ITGO.
 DR Interpro; IPR002064; DNA_pol_B.
 DR Interpro; IPR004578; Pol2.
 DR Pfam; PF00136; DNA_pol_B_1.
 DR Pfam; PF03104; DNA_pol_B_exo_1.
 DR PRINTS; PRO0106; DNAPOLB.
 DR SMART; SM00486; POLBc; 1.
 DR TIGRfam; TIGR00592; pol2; 1.
 DR PROSITE; PS00116; DNA POLYMERASE B; 1.
 DR Transferrase; DNA-directed DNA polymerase; DNA replication;
 KM DNA-binding
 SQ SEQUENCE 775 AA; 89781 MM; 71525D63C0C9CD9D CRC64;
 Query Match 28.6%; Score 1197.5; DB 1; Length 775;
 Best Local Similarity 36.1%; Pred. No. 4.3e-69;
 Matches 301; Conservative 139; Mismatches 296; Indels 97; Gaps 24;
 8 VLDSSEYVVGKEPOVYIIGIAENGERVVLIDRSFRPYFALLAPGA---DPKQV-AQRI 63
 2 IILDYITENGKPIIRVFK-KEGEFKIEYDRTFEFPYALLKDDSAIEDVKKYTAHSG 60
 64 ALSPPKSPDIIGVEDDKRYGPRPRVLRIRTVLPEAVREYRELKVNVDGVEDLEADIR 123
 61 TVAVKVR-----AEKVQKFLGRPIEWMKLYLEHPQDVPAIRAHRAVAVDIEYDIP 116
 124 AMRVLIDHDLFPFTWYRVEAPLEENKGFVVDKYLVKSRPEPLYGELAPTKLPDLRIL 183
 117 AKRVLIDKGLI-----PMEDG-----EELTML 138
 184 AFDIEVYVSKGSPRPERDPVIVAVKTDGDEVLFLAEGKDD-----RKPIREFV 233
 139 AVDIETLYHEGE-EFGTPIIMISYA--DGESEARVITWKKIDIPYDVVSTEREMIKRFL 195
 234 EYVRYVDPDIIVGNNHFPYLLRRARILGKLDVTRVGAEPITSVGH---VSVPG 290

```

Db 196 RVRREKDPVLITVNGDNFPAVYKKRCCELGKIKTLGRD-GSEPKIQRMGDFRANVKG 254
Qy 291 RANVPLDYAEEMPEIKTSLSEVAELTGMKSKSERVIINWWEIPDYWDPPKRRLLLO 350
Db 255 RIFEDLPVIRRTINLPITLLEAVEN-FGRKEKVAE--ETAOAMEGEGLEAVARY 311
Qy 351 ARDDVRATYGLAEKILPPALQSLVYTGLPIDOVGANSVGRLEWYLIRAFKXKELVPR 410
Db 312 SMDKAVTYELGREFPEMAQLSLLIGQSLMDVSSSTGNVEMFLLRKAYKKNELAPRK 371
Qy 411 VERPEET-----YRGATVLEPLRGVHENIAVLDFSSMTNIMIKYVGPDTLVRPGEKG 465
Db 372 PBERLARRRGGYAGGYKBERGLMDNLYLDFRSLYPSIITHTANSPTTLNRE----- 426
Qy 466 ECGGWE---APEVKHREKRCPPGFETVLERLLELRKVRAMKKYPPDSPEYRLDERO 522
Db 427 --GCKEYDVAPEVGHKFCDDPPGIPSLGDLLEEROKIKRKMKA-TVDPLEKKLDVRO 483
Qy 523 KALKVLANASYGWMGSGARVYCRCAKAVTAMGRHLRTAI-MIARKLGAKYIGDPTS 581
Db 484 RAIKILANSFYGYGYAKARVYCKECSASTAMGREYIEMTVIRELEKFGPKVLVADTNG 543
Qy 582 LFTY---DPE---KYENFKIKYBELG--PEIKLEKYKRLPFTAKKRYAGLLEDR 632
Db 544 LHATIPGADAEIVYKKAKKEPLKYNPKLPGLBLEIEGFRVGRFFV-TKKKAVIDEBCK 602
Qy 633 IDIVGFVARGDWCELAKEVOTKVEIVLKTSEVNKAVERKIVKLEEGKVPLEKVI 692
Db 603 ITRGLEIVRDMSEIALETQARVLEILHGDVEAVRIYKEVTEKLSKYEPPEKVI 662
Qy 693 WKTLSKLEEYTTAPHVVAKRLMSAGRVSPGDKIGVIVKSGGRISORAMPYFVMD 752
Db 663 HEGITRDLRDYKATGPVAAVAKRLAAGVIRPGVIGSIYIVKSGSIRGRLPADEF-D 721
Qy 753 PS--QIDVYVVDHGIIPALRLIGVFGITEKKLKAATGOKTLPDFLAKSK 803
Db 722 FTGHRVDAEYIENQVLEPAVERILKAGYRKEDLRYOKTKOVGLGAWLKVKGK 774

```

RESULT 9

DPOL METVO

STANDARD; PRT; 824 AA.

ID PS2025;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DB DNA polymerase (EC 2.7.7.7).

EN POL.

Methanococcus voltae.

Archaea; Euryarchaeota; Methanococci; Methanococcales;

Methanococcaceae; Methanococcus.

NCBI

NCBI

NCBI

NCBI

NCBI

NCBI

NCBI

NCBI

NCBI

NCBI

NCBI

NCBI

NCBI

NCBI

NCBI

NCBI

NCBI

NCBI

NCBI

NCBI

NCBI

NCBI

NCBI

NCBI

```

DR InterPro: IPR002064; DNA_pol_B.
DR InterPro: IPR004578; Pol2.
DR Pfam: PF00136; DNA_pol_B.1.
DR PRINTS: PR00106; DNA_POLB.
DR SMART: SM00486; POLB; 1.
DR TIGRFAMs: TIGR00592; pol2; 1.
DR PROSITE: PS00116; DNA_POLYMERASE_B_1.
DR Transferrase; DNA-directed DNA polymerase; DNA replication;
KW DNA-binding.
SQ
SEQUENCE 824 AA; 96754 MW; 094A630D8C1514FC CRC64;

```

Query Match

23.5%; Score 986; DB 1; Length 824;

Best Local Similarity 29.0%; Pred. 1.5e-55;

Matches 255; Conservative 159; Mismatches 288; Indels 178; Gaps 25;

```

Qy 9 LDSYEVNG-KEPOYIWMGANGERNVLIIDSPFPYFALLAGADPKQVAGRIALSR 67
Db 12 IDMYKCKLKEPEINLOKECE-----FRPYF--VDTSEPKIYIYDGLNO 57
Qy 68 -----PK-----SPITGEDDKRYFGRR 88
Db 58 EIDKLEPEFENNTSLKQVODLTINIEIKYISDYILNGKDISVSDFRNK--KERK 114
Qy 89 VLIRITLPEAVREYRELVRKVDGEDVLEADIRFAMRYLIDHDLFPFTYRVEAEPLN 148
Db 115 ICKVYKYRNHVKIIRIFYKEFG--KSYEFDPFLRRYIMDODIVSAKYS-EDNKKIN 170
Qy 149 KMGFRVDKYLVRSPREPLVYGBALAPTKLPDLRIIAFDIEVYSKQGSPPREDPVIVAV 208
Db 171 -----SIPELNCLIAFMELCYCK-EPNAKQDPIIMVNL 202
Qy 209 KTDGGEVL-----FIAGKDRKPIREFVEYVRYPPDIIVGNNHFDMPY 256
Db 203 FSKDYQKVIYTKKPFENSEYNGCVYVDEKELIQITIELKQY--DIVIYNGDNDFPY 260
Qy 257 LRRARILGIKLDVRRVGAPEPT-----SVHGIVSVPRGLAVDYDAEEMPEIKIK 309
Db 261 LKGRANIVIEIDPDMANSQOPOIKISKGIKNSKIPGIIHIDILPARKLNLVTR 320
Qy 310 STEFVAYLVGKKSEKRYIIMWEIPYWDPPKRRLLQVARDVATYGLAEKILPRA 369
Db 321 KLENVQELPKINKE--AVDYGIPRWE--TEDTTLKRAYEDALYTKMKNYFLPLE 375
Qy 370 IQLSVYTGLPIDOVGANSVGRLEWYLIRAFKXKELVPRV-----EPPEYTGALV 423
Db 376 IMFSRIVNOPLDYTSRMNSQWVEFLKRSFEGNNISPNRPSSSYREKAKSYEGYV 435
Qy 424 LEPRLGVHENIAVLDFSSMYENIWKYVGPDTLVRPGEKCGCWAEAVGHRRRCR 483
Db 436 REPLKGIQEDIVSLDFMILYPSILISHNISPEFVIY-----EKKER-ENNE 480
Qy 484 PGFFKTVLERLLELRKVRAMK---KYPDSPEYRLDEROKALKVLNANASYGWMG 539
Db 481 LGIIPKTLNELLSRRKHKMLKDKIOKNEFDE-EYSRLHEQGSIVKLANSHYGLAP 539
Qy 540 GAWYCRBACAKVNTAMGRHLIRTAIRNARKLGAKYIGDTGSLVTVDPKRV----- 592
Db 540 MARWYDCKAENVTLGKXYIQETIEKAEFEGKVIYADTGFPAKNDYDKLQKKEEN 599
Qy 593 -----NFKIKIEEL--GEPIKLEKVKR-LFTEAKKRYAAL 627
Db 600 DKSDKSNPKSKSEELIITKFKLKGINBELBEGHELEFGRFRGLFYT--KKRYALI 657
Qy 628 LEDGRIDIVGPAVARGDWCELAKEVOTKVEIVLKTSEVNKAVERKIVKLEEGKVP 687
Db 658 EDGSHLVVGLGVARRDMSNIAKQOAVIRLLEBGGVNLAKKIINKTIDNKKKNINIK 717
Qy 688 EKLIVKTLSEKLEEYTTAPHVVAKRLMSAGRVSPGDKIGVIVYKGGRIISORAMPY 747
Db 718 NDLLIHQTOLKNIIEYKSTAHIEVAKIKQGDGSRVAGVYSIYVKGSRISERA--- 774
Qy 748 FMYVDSQIDVYVVDHGIIPALRLIGVFGITEKKLKA 787

```

Db 775 ELLEYAGDYDINYIDNOVLPVIRIMESLGISEDELKNS 814

RESULT 10

PROL_PYRSD STANDARD; PRT; 1312 AA.

AC Q51334; Q51335; Q51336;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE DNA polymerase (EC 2.7.7.7) (Deep vent DNA polymerase) [Contains:
DNA polymerase PI-Pepi (EC 3.1.-.-) (Pep-GDS pol intein)].

GN POL.

OS Pyrococcus sp. (strain GB-D).

OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.

OX NCBI_TaxID=69013;

RP (1)

SEQUENCE FROM N.A., AND SEQUENCE OF 493-517.

MEDLINE=94094330; PubMed=8269515;
Xu M.-Q., Southworth M.W., Mersha F.B., Hornstra L.J., Peiter F.B.:
"In vitro protein splicing of purified precursor and the
identification of a branched intermediate.";
Cell 75:1371-1377(1993).

RL C611 75:1371-1377(1993).

CC -1- FUNCTION: IN ADDITION TO POLYMERASE ACTIVITY, THIS DNA POLYMERASE
EXHIBITS 3' TO 5' EXONUCLEASE ACTIVITY.

CC -1- FUNCTION: INTEIN ENCODED ENDONUCLEASES ARE THOUGHT TO MEDIATE
INTEIN MOBILITY BY SITE-SPECIFIC RECOMBINATION INITIATED BY
ENDONUCLEASE CLEAVAGE AT THE "HOMING SITE" IN GENE THAT LACK THE
INTEIN.

CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
+ {DNA} (N).

CC -1- PTM: THIS PROTEIN UNDERGOES A PROTEIN SELF SPLICING THAT INVOLVES
A POST-TRANSLATIONAL EXCISION OF THE INTERVENING REGION (INTEIN)
FOLLOWED BY PEPTIDE LIGATION.

CC -1- BIOTECHNOLOGY: Used in the PCR method because of its high
thermostability and low error rate. Sold by New England Biolabs.

CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY
-1- SIMILARITY: IN THE INTEIN SECTION; BELONGS TO THE HOMING
ENDONUCLEASE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.ebi.ac.uk/announce/>
or send an email to license@ebi.ac.uk).

CC -----

EMBL: U00707; AAA67130.1; -

EMBL: U00707; AAA67131.1; -

EMBL: U00707; AAA67132.1; -

DR HSSP; P56689; ITGO.

DR REBASE; 2619; PI-Pepi.

DR InterPro; IPR002064; DNA_pol_B.

DR InterPro; IPR003586; Hedgehog_hintc.

DR InterPro; IPR003587; Hedgehog_hintn.

DR InterPro; IPR002203; intein.

DR InterPro; IPR004042; intein_endonuc.

DR InterPro; IPR004578; Pol2.

DR Pfam; PF00136; DNA_pol_B; 1.

DR Pfam; PF03104; DNA_pol_B_exo; 1.

DR PRINTS; PR00106; DNAPOLB.

DR PRINTS; PR00379; INTEIN.

DR SMART; SM00305; Hintc; 1.

DR SMART; SM00306; Hintn; 1.

DR SMART; SM00486; POLBc; 1.

DR TIGRFAMs; TIGR00592; Pol2; 1.

DR PROSITE; PS00116; DNA_POLYMERASE_B; 1.

DR PROSITE; PS50818; INTEIN_C_TER; 1.

DR PROSITE; PS50819; INTEIN_ENDONUCLEASE; 1.

DR PROSITE; PS50817; INTEIN_N_TER; 1.

KW transferase; DNA-directed DNA polymerase; DNA replication;

KW DNA-binding; Hydrolase; Endonuclease; Autocatalytic cleavage;

KW Protein splicing; Inton homing;

FT CHAIN 1 492 DNA POLYMERASE, 1ST PART.

FT CHAIN 493 1029 ENDONUCLEASE PI-PSPI.

FT CHAIN 1030 1312 DNA POLYMERASE, 2ND PART.

SO SEQUENCE 1312 AA; 152852 MW; B62518805641D26A CRC64;

Query Match 23.1%; Score 965.5; DB 1; Length 1312;

Best Local Similarity 21.9%; Pred. No. 5.8e-54;

Matches 300; Conservative 154; Mismatches 282; Indels 635; Gaps 23;

Qy 8 VLDSYVEVKEPOVITIAENGERRVYLDISRPYFVYLLAPGADPKQV-----AQRIR 63

Db 2 ILDDADITTEGKPIIRIFK-KENGSEFVEYDRNRPYITALLKDDSDIDVRKITLIERHG 60

Qy 64 ALSRPSPIIGVEDDKRKRYGRRPRLVIRTVLPEAVREYRELKQVGDVEYLEADIRP 123

Db 61 KIVR-----IIDAERKRRKFLGRPLEVWRVLFEPHPDVPARIDKIREHSAVIDIFEVDIIF 116

Qy 124 AMRYLIDHDLFPPTWRVVEAPLENNKGRFVDKVLVSKRPEPLYGELAPTKLPDLRL 183

Db 117 AKRYLIDKGLI-----PMEGD-----EBIKL 138

Qy 184 AFDIEVYSKQSPRPERRPVIVAVKTDGDEVL-----PIASGKDRKPIREFVEY 235

Db 139 AFDIEITLHEGE-EFAKGPITIMSYADEEAKVITMKKIDLPVEYVSSREMIKRLKV 187

Qy 236 VKRVDPIIVGYNMNHEDWYLLRRALIGIKLDVTRRGAAPTTSVGH---VSYPGR 292

Db 198 IREKDPVITITNGDSFDLPYLVRRAKIGIKLPLRGD-GSEPRKQRLGDMTAVELIKRI 286

Qy 293 NDLYDAEMPEIKISLEVAEYLGWKKSRVLIIMMEIPYMDPKRPLLIQYAR 352

Db 257 HFDLYHIVIRRTIMLPYTLAEVYEA-IGKPKKQVVAH-EIEMAVETGKGLERVAKYSM 313

Qy 353 DDVATYGLAEKILPFAIOLSVYTGPLPDQVAMSGFRELWLLIAAFMKELPRVVE 412

Db 314 EDKATVYELREFFEPMEAOQLSRVGOPLDVSSSTGNLVEWLLKRAERNEELAPKPD 373

Qy 413 RPE-----ETRGATVLEPLRGVHENIAVLDFSSMYPNIMIKYVNPDTLVPRGEKGE 466

Db 374 EREYERLRARSYAGVYKEPEKGLMEGLVSLDRSLYPIIIHNVSPTLANE----- 427

Qy 467 CGGCE--AEVYKRRRCPPGPFKTVLERLRLKRVRAEMKPPDPSEYRLDEROK 523

Db 428 -GCREYDVAPEVGHKFCXDPGFIPLSLKRLDEROIKRKMKA-SKDPLEKTLDYROR 485

Qy 524 ALKVL----- 529

Db 486 ALKILANSILPEEWPLINKGVKIFRIGDFVDGLMKANQGVKKTGDTEVLEAGIHAF 545

Qy 530 ----- 529

Db 546 SPDRKSKARVMAKVAIRHRYSGVYRIYVNSGRKITTIEGHSLPYRNGDLVEATGED 605

Qy 530 ----- 529

Db 606 VKIGDLAVRSVNLPEKERLNIIVELLNLSPETEDILITIPVKGRKNFPKGMRLTLR 665

Qy 530 ----- 529

Db 666 WIFGEKRVATSRVYLHLENLGYIRLKIYDIIDKEGLEKRYLYEKLVDVRYVNGNK 725

Qy 530 ----- 529

Db 726 RYLVEFNARVDYISLPEBELKEMRIGTNGFRMGTFTVIDDPAKLGLGYVSEGSARK 785

Qy 530 ----- 529

Db 786 WKQGTGWSYTVLYNENDEVLDMEHLAKKPFQKVRGKGVVEIPKMAVITFESLGT 845

Qy 530 ----- 529

Db 846 LAENKAVPEVITFTSSKGVNAPLEGYTIGGDVHPKRVLTSTSELLNGLVLLNSIG 905
 QY 530 ----- 529
 Db 906 VSAIKGDSGVYRVYNEELKFTYRKKKNVSHIVPKDILKETFGKVFQKNISYKPF 965
 QY 530 ----- 529
 Db 966 RELVNGKLDREKAKRIEMLNGDIVLDRVVEIKREYDGYVDSYDEBNFLAGFGL 1025
 QY 530 ---NASYGVMGSGARWYCRBCAATVAMGRHLIR-TAINIRKLGKLVYIGDTSLEFY 585
 Db 1026 YAHNSYGYGYAKARWYCKEACASVTAMGREYIEFVRKELEPKGPFVCLIDTDLVAT 1085
 QY 586 Y---DEKVE---NFIKIKEELG--FEIKLEKVKRLEFTEAKKRYAGLLEBGRIDIV 636
 Db 1086 IPGAKPEIKKKALFEVDYINAKLPGELLEEGFVYRGFFV-TKKYALIDEGKILIR 1144
 QY 637 GFEAVRGDWCELAKEVOKKVEIYLAKESEVNAKVEYKIVALEBGRVPEKLYWKTL 696
 Db 1145 GLEIVKRDMSIAKETQAKVAILKGNVEANVKLVKVTIKUSKEIPEKCLVYEQI 1204
 QY 697 SKLEETTEAPHYVAAKRMLSAGTVSPQDKIGYIVYKGGRIISQRAPEYMWKD--- 752
 Db 1205 TRPHEKAIQPHVAVARLARAVKVRPMVIGYIVLKGDPISKRA---ILAEFDR 1261
 QY 753 PSQIDVTVYVDHQIIPALRILGYFGITEKCLKASATVQCKTLFDFLAKSK 803
 Db 1262 KHKYDAVYIENQVLPVIRILBAFGYRKEDLRWKQTKQTLTAMLNKKK 1312

RESULT 11
 ID POL PYRHO STANDARD, PRT; 1235 AA.
 AC 059610;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE DNA polymerase (EC 2.7.7) (Contains: Plo pol intein (Plo pol I
 intein))
 GN POL OR PH1947 OR PH87047.
 OS Pyrococcus horikoshii.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 Pyrococcus.
 KX NCBI_TaxID=53953;
 RN (1)
 SEQUENCE FROM N.A.

STRAIN=OT3;
 MEDLINE=98344137; PubMed=9679194;
 KAWABAYASHI Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
 Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hongoyama A., Nagai Y.,
 Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
 Funahashi T., Tanaka T., Kudo H., Yamazaki J., Kushiida N., Oguchi A.,
 Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
 Maehuri Y., Shizuya H., Kikuchi H.;
 "Complete sequence and gene organization of the genome of a hyper-
 thermophilic archaeobacterium, *Pyrococcus horikoshii* OT3";
 RNA Res. 5:55-76(1998).

RT DNA Res. 5:55-76(1998).
 -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
 + (DNA) (N)
 -1- P.TM. THIS PROTEIN UNDERGOES A PROTEIN SELF SPLICING THAT INVOLVES
 A POST-TRANSCRIPTIONAL EXCISION OF THE INTERVENING REGION (INTEIN)
 FOLLOWED BY PEPTIDE LIGATION (POTENTIAL).
 -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL, AP000007, BAA31074.1; --
 DR HSSP, P56689, ITGO.
 DR InterPro: IPR002064, DNA_pol_B.
 DR InterPro: IPR003586, Hedgehog_hinc.
 DR InterPro: IPR003587, Hedgehog_hinc.
 DR InterPro: IPR002203, Intein.
 DR InterPro: IPR004042, Intein_endonuc.
 DR InterPro: IPR004578, Pol2.
 DR Pfam: PF00136, DNA_pol_B_exo_1.
 DR Pfam: PF03104, DNA_pol_B_exo_1.
 DR PRINTS: PR00106, DNAPOLB.
 DR SMART: SM00305, HincC_1.
 DR SMART: SM00306, HincN_1.
 DR SMART: SM00486, PolBc_1.
 DR TIGRfam: TIGR00592, pol2_1.
 DR PROSITE, PS00116, DNA_POLYMERASE_B_1.
 DR PROSITE, PS00818, INTEIN_CTER_1.
 DR PROSITE, PS00819, INTEIN_ENDONUCLEASE_1.
 DR PROSITE, PS00817, INTEIN_NTER_1.
 DR TRANSFERASE: DNA-directed DNA polymerase; DNA replication;
 KM DNA-binding; Autocatalytic cleavage; Protein splicing;
 KM Complete proteome.
 FT CHAIN 1 492 DNA POLYMERASE, 1ST PART (POTENTIAL).
 FT CHAIN 493 952 PHO POL INTEIN (POTENTIAL).
 FT CHAIN 953 1235 DNA POLYMERASE, 2ND PART (POTENTIAL).
 SQ SEQUENCE 1235 AA; 143086 MM; 73CCTAA14873CCE4 CRC64;

Query Match 22.8%; Score 955.5; DB 1; Length 1235;
 Best Local Similarity 23.1%; Pred. No. 2.3e-53;
 Matches 299; Conservative 145; Mismatches 291; Indels 559; Gaps 24;

QY 8 VLDSSEYVNGKEREPOVIWGIANGERVYLIDSRFRPYFALL---APGADPQVQVR 63
 DB 2 ILADADYTDGKPIIRIK-KENGEFVNEYDRFRPYVYALLRDSALIDEIKITQNRG 60
 QY 64 ALSRPSPIIGVEDDKRYRPRRVRVIRITVPEANREYELVKNVGDVEDLVENDIF 123
 DB 61 KVRV---IVETEKQKFLDRPLEVWKLYLHPQDVPAIRKREHPAVADVFEYDIF 116
 QY 124 AMRYLIDHDLFPFTVRYVEAEPLLENKGFVVDKVLVNSRPEPLYGBALPTKLPDLRL 183
 DB 117 AKRYLIDKGL-----TPMEGN-----EKLTPLE 138
 QY 184 AFDIEVYSKQSPREPRDPVIVIAVKTDDGDEVL-----FLAEGDKRKEIFREVEY 235
 DB 139 AVDIEFLYHGE-EPGKGPVIMISYADSECAKITWKKIDLPYEVVSSREMIKRLIRV 197
 QY 236 VKRYDPIIYGVNNHNPMPYLLRRARILGIKLDVTRRVGAEPPTSVHGH---VSVFGR 292
 DB 198 IKEKDPVLIITNGNDFPFLRLRAKGIKTL-LLGHRNNSPPMKQKGSGLAVEIKGRI 256
 QY 293 NVLDYVDAEMPEIKSLSEVAEYGVMMKSRVYINMMELPYVMDPKKRLLQVYR 352
 DB 257 HFDLPYIRKIRTNLPYTLLENVYAL-FGKPKENYAD-EIAAKMETGSGLERVAKTSM 313
 QY 353 DVRYATYGAELKILPAILQSLSYVGLPFDQYGMASVGRLELYILRAAFKKLVNRYE 412
 DB 314 EDARVTEYEGREFPEBAQLARLVGPVMDVSRSGNLVEPELRKAYRENEELANPKD 373
 QY 413 RPE-----ETRYGALVLEPLRGVHENIAVLDFSSMYPNIMKYNVGPDTLVPRGKCE 466
 DB 374 EKEERRLRESYEGYVPEPEKGLMEGIVSLDFRSIYSIITINVSPTLNE----- 427
 QY 467 CGCWE---APEYVHFRRCPPGFYKTYLLETLKRYVALEMKYPPDSPEYVLLDEROK 523
 DB 428 -GCSEYDAVPKGFHFCDFKDPGFIPLGLLEEROKIKKKWE-SKDPVEKKLLDYRQR 485
 QY 524 ALKTLA----- 529
 DB 486 ARIKLANSILPDEWLPYVENKFAVYKIGFDIDREIEENARVYKDDETELLEVOKLAL 545
 QY 530 ----- 529

Db 546 SPNETKSEELKVKVALLIRHRYSGKVSILKSGRRITKTSGHSLFSGNKLKVRGDE 605
 QY 530 ----- 529
 Db 606 LKPDLLVVPGRLLKPSKQVLANIVELLKLPBEETSINVMIPDKGRKQFPKMLTLY 665
 QY 530 ----- 529
 Db 666 WIFEGSRPRTAGRYLKHLERLGVYVYKLRGCEVLDWESLKRKYLYETLLKNLKNVNS 725
 QY 530 ----- 529
 Db 726 RAYWERNSLRDVVSILMPBELKEMIIIGEPGRKIGTIFDVDSFAKLLGYIISGDEVK 785
 QY 530 ----- 529
 Db 786 DRVYFHSKDQNVLEDAKLAELKFGVRRGRGYLEVSGKISHALFRVLAEGKRIPFT 845
 QY 530 ----- 529
 Db 846 SPMDIKVAFKLGANGNAELFTSTKSELVNOJLLNLSIGVSDIKIEHEKGYRVYINK 905
 QY 530 ----- 529
 Db 906 KESNGDIVLDSVESIEVEKEYGYVDLSDVENENFLVGFGLYAHNSYGYGAKARW 965
 QY 544 YCRCAKAVTAMGRHLIRTAIINARKGLKRYIGDTSLEFTY----DPEKVE---NFI 595
 Db 966 YCKEASAVTAMGRQYIDLVRRELBARGFKVLIDTDLVYTTIGVYKDWEMVKRRALFEV 1025
 QY 596 KIIKEELG--FEIKLEKRYKLFTEAKRYAGLLEDRIDIVGFEAVRGDMCELAKEVQ 653
 Db 1026 DYNISKLPVLELEVEGYAGFV--TKKVALIDBEKLYTRLEIVYRMSIAKETQ 1084
 QY 654 TKVVEIVLKTSEVNAKAVEYKRIKVELEEGVPIEKLYIMTSLKRLSEYTTAPHVAA 713
 Db 1085 ARVLEALIKHGNVEAVKIVADVTEKLTNYVEPEKLTIVYQITRPINEYKALIPHVAA 1144
 QY 714 KRMSAGRYVSPGDKIGVYIKKSGRISORAMPYFMK--DP--SQDVTYVVOHIIIPA 769
 Db 1145 KRLWARGIKVAPGVNIGIVIRGDSPIKRA--ISIEFPRKHKDAEYIENQVIPA 1201
 QY 770 ALRILGYFITEKLLKASATQKTLFDFL-AKKS 802
 Db 1202 VERILKAFGYREDLPRMOKTKQVGLGAMIKYKKS 1235

SQLT 12
 L-METTH
 AC 027276; STANDARD; PRT; 7586 AA.
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE DNA polymerase (EC 2.7.7.7).
 GN POL OR POLB1 OR MTH1208.
 OS Methanobacterium thermoautotrophicum.
 OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
 OC Methanobacteriaceae; Methanothermobacter.
 OX NCBI_TaxID=187420;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Delta H.
 RX MEDLINE=803751; PubMed=9371463.
 RA Smith D.R., Doucette-Stamm L.A., DeLoughery C., Lee H.-M., Dubois J.,
 Aldege T., Baahirzadeh R., Blakely D., Cook R., Gilbert K.,
 Harrison D., Hoang L., Keagle P., Lumm W., Pochler B., Qiu D.,
 Spadator R., Viscare R., Wang Y., Wierzbowski J., Gibson R.,
 Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
 Madougall S., Shimer G., Goyal A., Pietrovski S., Church G.M.,
 Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.,
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum
 deltaH: functional analysis and comparative genomics.";

RL J. Bacteriol. 179:7135-7155(1997).
 CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
 CC +[DNA](N).
 CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: AE000889; AAB85697.1; -
 DR HSSP: P56689; ITGO.
 DR InterPro: IPR002064; DNA_pol_B.
 DR Pfam: PF00136; DNA_pol_B_1.
 DR Pfam: PF03104; DNA_pol_B_exo; 1.
 DR PRINTS: PR00106; DNAPOLB.
 DR SMART: SM00486; POLB; 1.
 DR PROSITE: PS00116; DNA_POLYMERASE_B; 1.
 KW Transferrase; DNA-directed DNA polymerase; DNA replication;
 KW DNA-binding; Complete proteome.
 SQ SEQUENCE 586 AA; 67965 MW; D406B5BC399B951A CRC64;

Query Match 21.8%; Score 915; DB 1; Length 586;
 Best Local Similarity 35.4%; Pred. No. 3, 4e-51;
 Matches 218; Conservative 110; Mismatches 226; Indels 62; Gaps 12;

QY 1 MTEVFTVLDSSYEYVKEPVIIMGIAENG--ERVVLIDRSRPPEYVALLAGADPKOV 58
 Db 1 MEYRWVLDIDVYVDEVPVIRLFGKSGSGNEPIIADRSRPPIYAL-----PTDL 54
 QY 59 AGRIRALSRKSPDIIGVEDDKRYFGRRPVLRIRTLPEAVREYELVANGVDEVLE 118
 Db 55 DECLRELELELEKLEWK--MEDLRPTEVIRIEPRHPQDVKIRIDRLDESVDIRE 112
 QY 119 ADIRFARMYLIHDLFP-----FTYRVAEAPLEMKSGRVKRYLYVKSREPYLEALA 173
 Db 113 HDIPFRRYRILDKSIVPMEELEFGVVEVDSAPSVTITDVTEVTVGRVOSTGSGAHG---- 168
 QY 174 PTKLPDLRIAPFIEVYVSKSGSPRPDPVIVYAVK-----TDDGDEVLFIAEGKD 224
 Db 169 -----LDLISFIEVRANPFGMPDPKDELIMGVAGNMGESVISTAGDHLDFEVVED 222
 QY 225 DRKPIREFEVYKRYDPIIIVGYNHNFPMVYLLRRARILGILDY-----TRVGA 276
 Db 223 ERELLEFPAIVIDKKPDILVGNDSNDFPYITRAAILGAEILDGWDGSKIRTVRRGF 282
 QY 277 EPTTSVHGHSVSGRLANDVLYDAEMPEIKISLEVAEYLCWKKSERVIINMWEIPD 336
 Db 283 ANATAIKGV-----HVDLYPVMRRYMLDRYTLERVYQELGEBEKIDLPGRDWE-- 333
 QY 337 YMDPKRPLLLQYARDVATYGLAEKILPFAIOLSVYTLGFLPDVGAMSVGFRLEWY 396
 Db 334 YMRDELRLDELFRYSDDVATIRIAEKILPLNLEIRLVGQPLFPISRVATQOQHEWFL 393
 QY 397 IRAAFKKEKLVNRRVERPEETV-----GAVLEPLRGVHENIIVLDSSMYPNMIKY 450
 Db 394 VRKAYQYGLVLPKPSQSDPSRRGRRAVGVYKEPEKGMENIYQFDPRLSLPSIISK 453
 QY 451 NVGPDITLVPRGECGCGCEAPVHRRRCPPGPFKYVLELRLKRVRAEMKKYPP 510
 Db 454 NISPDITLTDDEE-----ECVVAPEYGYRFPKSPRGVPSVIGILBERVAKIKEMKG-SD 508
 QY 511 DSEYRLDEROKALKVLANASGYGWSGARYCECAKAVTAMGRHLIRTAIINARKL 570
 Db 509 DPEERKILVQALKRLANTMTGYGYGSRFRYSMECAEATAMGRDIYIKTITAEER 568
 QY 571 GLKVIYGDTSLEFTY 586
 Db 569 GFHTVYADTGGFYATY 584

RESULT 13
 ID POLYMERASE STANDARD, PRT, 1523 AA.
 AC P74918;
 DT 15-DEC-1998 (rel. 37, Created)
 DT 15-DEC-1998 (rel. 37, Last sequence update)
 DT 15-JUN-2002 (rel. 41, Last annotation update)
 DE DNA polymerase (EC 2.7.7.7) [Pol Tfu] [Contains: Endonuclease PI-Tfu
 (EC 3.1.-.-) (Tfu pol-1 intein); Endonuclease PI-Tfu (EC 3.1.-.-)
 (Tfu pol-2 intein)].
 GN POL.
 OS Thermococcus fumicolans.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Thermococcus.
 NCBI_Taxid=46540;
 [1]
 SEQUENCE FROM N.A.
 STRAIN=ST557;
 Cambo M., Querellou J.;
 Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
 [2]
 RP CHARACTERIZATION OF INTEINS.
 RC STRAIN=ST557;
 RA MEDLINE=20112788; PubMed=10644683;
 RA Saveri I., Ozanne V., Dietrich J., Masson J.-M.;
 RT "Inteins of Thermococcus fumicolans DNA polymerase are endonucleases
 with distinct enzymatic behaviors.";
 RL J. Biol. Chem. 275:2335-2341 (2000).
 CC -1- FUNCTION: PI-Tfu recognizes and cleaves a minimal sequence of 16
 base pairs (bp) on supercoiled DNA with either Mn(2+) or Mg(2+) as
 cofactor. It cleaves linear DNA only with Mn(2+) and requires a
 19-bp minimal recognition sequence. The optimal temperature for
 activity is 70 degrees Celsius.
 CC -1- FUNCTION: PI-Tfu is a highly active homing endonuclease using
 Mg(2+) as cofactor. Its minimal recognition and cleavage site is
 21 bp long either on linear or circular DNA substrates. Its
 endonuclease activity is strongly inhibited by the 3' digestion
 product, which remains bound to the enzyme after the cleavage
 reaction. The optimal temperature for activity is 70 degrees
 Celsius.
 CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
 + [DNA] (N).
 CC -1- PTM: THIS PROTEIN UNDERGOES A PROTEIN SELF SPLICING THAT INVOLVES
 A POST-TRANSLATIONAL EXCISION OF THE TWO INTERVENING REGION
 (INTEINS) FOLLOWED BY PEPTIDE LIGATION
 CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.
 CC -1- SIMILARITY: IN THE INTEIN SECTION; BELONGS TO THE HOMING
 ENDONUCLEASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z69882; CA93738.1; -
 DR HSSP; P56689; 1TGO.
 DR REBASE; 4500; PI-Tfu1.
 DR InterPro; IPRO02064; DNA_pol_B.
 DR InterPro; IPRO03586; Hcdgheog_hlnc.
 DR InterPro; IPRO03587; Hcdgheog_hlnc.
 DR InterPro; IPRO02203; intein.
 DR InterPro; IPRO04578; intein_endonuc.
 DR Pfam; PF00136; DNA_pol_B; 2.
 DR Pfam; PF03104; DNA_pol_B_exo; 1.
 DR PRINTS; PR00379; INTEIN.
 DR SMART; SM00305; Hlnc; 2.
 DR SMART; SM00306; Hlnc; 2.

DR SMART; SM00486; POLB; 1.
 DR TIGRfams; TIGR00592; pol2; 1.
 DR PROSITE; PS00116; DNA POLYMERASE B; FALSE_NEG.
 DR PROSITE; PS50818; INTEIN C TER; 2.
 DR PROSITE; PS50819; INTEIN ENDONUCLEASE; 2.
 DR PROSITE; PS50817; INTEIN N TER; 2.
 KW transferase; DNA-directed DNA polymerase; DNA replication;
 KM DNA-binding; Hydrolyase; Endonuclease; Autocatalytic cleavage;
 KM Protein splicing; Intron homing; Magnesium; Manganese;
 FT CHAIN 1 406
 FT CHAIN 2 766
 FT CHAIN 3 900
 FT CHAIN 4 901
 FT CHAIN 5 1282
 FT CHAIN 6 1283
 FT CHAIN 7 1523
 FT CHAIN 8 1523
 FT CHAIN 9 1523
 FT CHAIN 10 1523
 FT CHAIN 11 1523
 FT CHAIN 12 1523
 FT CHAIN 13 1523
 FT CHAIN 14 1523
 FT CHAIN 15 1523
 FT CHAIN 16 1523
 FT CHAIN 17 1523
 FT CHAIN 18 1523
 FT CHAIN 19 1523
 FT CHAIN 20 1523
 FT CHAIN 21 1523
 FT CHAIN 22 1523
 FT CHAIN 23 1523
 FT CHAIN 24 1523
 FT CHAIN 25 1523
 FT CHAIN 26 1523
 FT CHAIN 27 1523
 FT CHAIN 28 1523
 FT CHAIN 29 1523
 FT CHAIN 30 1523
 FT CHAIN 31 1523
 FT CHAIN 32 1523
 FT CHAIN 33 1523
 FT CHAIN 34 1523
 FT CHAIN 35 1523
 FT CHAIN 36 1523
 FT CHAIN 37 1523
 FT CHAIN 38 1523
 FT CHAIN 39 1523
 FT CHAIN 40 1523
 FT CHAIN 41 1523
 FT CHAIN 42 1523
 FT CHAIN 43 1523
 FT CHAIN 44 1523
 FT CHAIN 45 1523
 FT CHAIN 46 1523
 FT CHAIN 47 1523
 FT CHAIN 48 1523
 FT CHAIN 49 1523
 FT CHAIN 50 1523
 FT CHAIN 51 1523
 FT CHAIN 52 1523
 FT CHAIN 53 1523
 FT CHAIN 54 1523
 FT CHAIN 55 1523
 FT CHAIN 56 1523
 FT CHAIN 57 1523
 FT CHAIN 58 1523
 FT CHAIN 59 1523
 FT CHAIN 60 1523
 FT CHAIN 61 1523
 FT CHAIN 62 1523
 FT CHAIN 63 1523
 FT CHAIN 64 1523
 FT CHAIN 65 1523
 FT CHAIN 66 1523
 FT CHAIN 67 1523
 FT CHAIN 68 1523
 FT CHAIN 69 1523
 FT CHAIN 70 1523
 FT CHAIN 71 1523
 FT CHAIN 72 1523
 FT CHAIN 73 1523
 FT CHAIN 74 1523
 FT CHAIN 75 1523
 FT CHAIN 76 1523
 FT CHAIN 77 1523
 FT CHAIN 78 1523
 FT CHAIN 79 1523
 FT CHAIN 80 1523
 FT CHAIN 81 1523
 FT CHAIN 82 1523
 FT CHAIN 83 1523
 FT CHAIN 84 1523
 FT CHAIN 85 1523
 FT CHAIN 86 1523
 FT CHAIN 87 1523
 FT CHAIN 88 1523
 FT CHAIN 89 1523
 FT CHAIN 90 1523
 FT CHAIN 91 1523
 FT CHAIN 92 1523
 FT CHAIN 93 1523
 FT CHAIN 94 1523
 FT CHAIN 95 1523
 FT CHAIN 96 1523
 FT CHAIN 97 1523
 FT CHAIN 98 1523
 FT CHAIN 99 1523
 FT CHAIN 100 1523
 FT CHAIN 101 1523
 FT CHAIN 102 1523
 FT CHAIN 103 1523
 FT CHAIN 104 1523
 FT CHAIN 105 1523
 FT CHAIN 106 1523
 FT CHAIN 107 1523
 FT CHAIN 108 1523
 FT CHAIN 109 1523
 FT CHAIN 110 1523
 FT CHAIN 111 1523
 FT CHAIN 112 1523
 FT CHAIN 113 1523
 FT CHAIN 114 1523
 FT CHAIN 115 1523
 FT CHAIN 116 1523
 FT CHAIN 117 1523
 FT CHAIN 118 1523
 FT CHAIN 119 1523
 FT CHAIN 120 1523
 FT CHAIN 121 1523
 FT CHAIN 122 1523
 FT CHAIN 123 1523
 FT CHAIN 124 1523
 FT CHAIN 125 1523
 FT CHAIN 126 1523
 FT CHAIN 127 1523
 FT CHAIN 128 1523
 FT CHAIN 129 1523
 FT CHAIN 130 1523
 FT CHAIN 131 1523
 FT CHAIN 132 1523
 FT CHAIN 133 1523
 FT CHAIN 134 1523
 FT CHAIN 135 1523
 FT CHAIN 136 1523
 FT CHAIN 137 1523
 FT CHAIN 138 1523
 FT CHAIN 139 1523
 FT CHAIN 140 1523
 FT CHAIN 141 1523
 FT CHAIN 142 1523
 FT CHAIN 143 1523
 FT CHAIN 144 1523
 FT CHAIN 145 1523
 FT CHAIN 146 1523
 FT CHAIN 147 1523
 FT CHAIN 148 1523
 FT CHAIN 149 1523
 FT CHAIN 150 1523
 FT CHAIN 151 1523
 FT CHAIN 152 1523
 FT CHAIN 153 1523
 FT CHAIN 154 1523
 FT CHAIN 155 1523
 FT CHAIN 156 1523
 FT CHAIN 157 1523
 FT CHAIN 158 1523
 FT CHAIN 159 1523
 FT CHAIN 160 1523
 FT CHAIN 161 1523
 FT CHAIN 162 1523
 FT CHAIN 163 1523
 FT CHAIN 164 1523
 FT CHAIN 165 1523
 FT CHAIN 166 1523
 FT CHAIN 167 1523
 FT CHAIN 168 1523
 FT CHAIN 169 1523
 FT CHAIN 170 1523
 FT CHAIN 171 1523
 FT CHAIN 172 1523
 FT CHAIN 173 1523
 FT CHAIN 174 1523
 FT CHAIN 175 1523
 FT CHAIN 176 1523
 FT CHAIN 177 1523
 FT CHAIN 178 1523
 FT CHAIN 179 1523
 FT CHAIN 180 1523
 FT CHAIN 181 1523
 FT CHAIN 182 1523
 FT CHAIN 183 1523
 FT CHAIN 184 1523
 FT CHAIN 185 1523
 FT CHAIN 186 1523
 FT CHAIN 187 1523
 FT CHAIN 188 1523
 FT CHAIN 189 1523
 FT CHAIN 190 1523
 FT CHAIN 191 1523
 FT CHAIN 192 1523
 FT CHAIN 193 1523
 FT CHAIN 194 1523
 FT CHAIN 195 1523
 FT CHAIN 196 1523
 FT CHAIN 197 1523
 FT CHAIN 198 1523
 FT CHAIN 199 1523
 FT CHAIN 200 1523
 FT CHAIN 201 1523
 FT CHAIN 202 1523
 FT CHAIN 203 1523
 FT CHAIN 204 1523
 FT CHAIN 205 1523
 FT CHAIN 206 1523
 FT CHAIN 207 1523
 FT CHAIN 208 1523
 FT CHAIN 209 1523
 FT CHAIN 210 1523
 FT CHAIN 211 1523
 FT CHAIN 212 1523
 FT CHAIN 213 1523
 FT CHAIN 214 1523
 FT CHAIN 215 1523
 FT CHAIN 216 1523
 FT CHAIN 217 1523
 FT CHAIN 218 1523
 FT CHAIN 219 1523
 FT CHAIN 220 1523
 FT CHAIN 221 1523
 FT CHAIN 222 1523
 FT CHAIN 223 1523
 FT CHAIN 224 1523
 FT CHAIN 225 1523
 FT CHAIN 226 1523
 FT CHAIN 227 1523
 FT CHAIN 228 1523
 FT CHAIN 229 1523
 FT CHAIN 230 1523
 FT CHAIN 231 1523
 FT CHAIN 232 1523
 FT CHAIN 233 1523
 FT CHAIN 234 1523
 FT CHAIN 235 1523
 FT CHAIN 236 1523
 FT CHAIN 237 1523
 FT CHAIN 238 1523
 FT CHAIN 239 1523
 FT CHAIN 240 1523
 FT CHAIN 241 1523
 FT CHAIN 242 1523
 FT CHAIN 243 1523
 FT CHAIN 244 1523
 FT CHAIN 245 1523
 FT CHAIN 246 1523
 FT CHAIN 247 1523
 FT CHAIN 248 1523
 FT CHAIN 249 1523
 FT CHAIN 250 1523
 FT CHAIN 251 1523
 FT CHAIN 252 1523
 FT CHAIN 253 1523
 FT CHAIN 254 1523
 FT CHAIN 255 1523
 FT CHAIN 256 1523
 FT CHAIN 257 1523
 FT CHAIN 258 1523
 FT CHAIN 259 1523
 FT CHAIN 260 1523
 FT CHAIN 261 1523
 FT CHAIN 262 1523
 FT CHAIN 263 1523
 FT CHAIN 264 1523
 FT CHAIN 265 1523
 FT CHAIN 266 1523
 FT CHAIN 267 1523
 FT CHAIN 268 1523
 FT CHAIN 269 1523
 FT CHAIN 270 1523
 FT CHAIN 271 1523
 FT CHAIN 272 1523
 FT CHAIN 273 1523
 FT CHAIN 274 1523
 FT CHAIN 275 1523
 FT CHAIN 276 1523
 FT CHAIN 277 1523
 FT CHAIN 278 1523
 FT CHAIN 279 1523
 FT CHAIN 280 1523
 FT CHAIN 281 1523
 FT CHAIN 282 1523
 FT CHAIN 283 1523
 FT CHAIN 284 1523
 FT CHAIN 285 1523
 FT CHAIN 286 1523
 FT CHAIN 287 1523
 FT CHAIN 288 1523
 FT CHAIN 289 1523
 FT CHAIN 290 1523
 FT CHAIN 291 1523
 FT CHAIN 292 1523
 FT CHAIN 293 1523
 FT CHAIN 294 1523
 FT CHAIN 295 1523
 FT CHAIN 296 1523
 FT CHAIN 297 1523
 FT CHAIN 298 1523
 FT CHAIN 299 1523
 FT CHAIN 300 1523
 FT CHAIN 301 1523
 FT CHAIN 302 1523
 FT CHAIN 303 1523
 FT CHAIN 304 1523
 FT CHAIN 305 1523
 FT CHAIN 306 1523
 FT CHAIN 307 1523
 FT CHAIN 308 1523
 FT CHAIN 309 1523
 FT CHAIN 310 1523
 FT CHAIN 311 1523
 FT CHAIN 312 1523
 FT CHAIN 313 1523
 FT CHAIN 314 1523
 FT CHAIN 315 1523
 FT CHAIN 316 1523
 FT CHAIN 317 1523
 FT CHAIN 318 1523
 FT CHAIN 319 1523
 FT CHAIN 320 1523
 FT CHAIN 321 1523
 FT CHAIN 322 1523
 FT CHAIN 323 1523
 FT CHAIN 324 1523
 FT CHAIN 325 1523
 FT CHAIN 326 1523
 FT CHAIN 327 1523
 FT CHAIN 328 1523
 FT CHAIN 329 1523
 FT CHAIN 330 1523
 FT CHAIN 331 1523
 FT CHAIN 332 1523
 FT CHAIN 333 1523
 FT CHAIN 334 1523
 FT CHAIN 335 1523
 FT CHAIN 336 1523
 FT CHAIN 337 1523
 FT CHAIN 338 1523
 FT CHAIN 339 1523
 FT CHAIN 340 1523
 FT CHAIN 341 1523
 FT CHAIN 342 1523
 FT CHAIN 343 1523
 FT CHAIN 344 1523
 FT CHAIN 345 1523
 FT CHAIN 346 1523
 FT CHAIN 347 1523
 FT CHAIN 348 1523
 FT CHAIN 349 1523
 FT CHAIN 350 1523
 FT CHAIN 351 1523
 FT CHAIN 352 1523
 FT CHAIN 353 1523
 FT CHAIN 354 1523
 FT CHAIN 355 1523
 FT CHAIN 356 1523
 FT CHAIN 357 1523
 FT CHAIN 358 1523
 FT CHAIN 359 1523
 FT CHAIN 360 1523
 FT CHAIN 361 1523
 FT CHAIN 362 1523
 FT CHAIN 363 1523
 FT CHAIN 364 1523
 FT CHAIN 365 1523
 FT CHAIN 366 1523
 FT CHAIN 367 1523
 FT CHAIN 368 1523
 FT CHAIN 369 1523
 FT CHAIN 370 1523
 FT CHAIN 371 1523
 FT CHAIN 372 1523
 FT CHAIN 373 1523
 FT CHAIN 374 1523
 FT CHAIN 375 1523
 FT CHAIN 376 1523
 FT CHAIN 377 1523
 FT CHAIN 378 1523
 FT CHAIN 379 1523
 FT CHAIN 380 1523
 FT CHAIN 381 1523
 FT CHAIN 382 1523
 FT CHAIN 383 1523
 FT CHAIN 384 1523
 FT CHAIN 385 1523
 FT CHAIN 386 1523
 FT CHAIN 387 1523
 FT CHAIN 388 1523
 FT CHAIN 389 1523
 FT CHAIN 390 1523
 FT CHAIN 391 1523
 FT CHAIN 392 1523
 FT CHAIN 393 1523
 FT CHAIN 394 1523
 FT CHAIN 395 1523
 FT CHAIN 396 1523
 FT CHAIN 397 1523
 FT CHAIN 398 1523
 FT CHAIN 399 1523
 FT CHAIN 400 1523
 FT CHAIN 401 1523
 FT CHAIN 402 1523
 FT CHAIN 403 1523
 FT CHAIN 404 1523
 FT CHAIN 405 1523
 FT CHAIN 406 1523
 FT CHAIN 407 1523
 FT CHAIN 408 1523
 FT CHAIN 409 1523
 FT CHAIN 410 1523
 FT CHAIN 411 1523
 FT CHAIN 412 1523
 FT CHAIN 413 1523
 FT CHAIN 414 1523
 FT CHAIN 415 1523
 FT CHAIN 416 1523
 FT CHAIN 417 1523
 FT CHAIN 418 1523
 FT CHAIN 419 1523
 FT CHAIN 420 1523
 FT CHAIN 421 1523
 FT CHAIN 422 1523
 FT CHAIN 423 1523
 FT CHAIN 424 1523
 FT CHAIN 425 1523
 FT CHAIN 426 1523
 FT CHAIN 427 1523
 FT CHAIN 428 1523
 FT CHAIN 429 1523
 FT CHAIN 430 1523
 FT CHAIN 431 1523
 FT CHAIN 432 1523
 FT CHAIN 433 1523
 FT CHAIN 434 1523
 FT CHAIN 435 1523
 FT CHAIN 436 1523
 FT CHAIN 437 1523
 FT CHAIN 438 1523
 FT CHAIN 439 1523
 FT CHAIN 440 1523
 FT CHAIN 441 1523
 FT CHAIN 442 1523
 FT CHAIN 443 1523
 FT CHAIN 444 1523
 FT CHAIN 445 1523
 FT CHAIN 446 1523
 FT CHAIN 447 1523
 FT CHAIN 448 1523
 FT CHAIN 449 1523
 FT CHAIN 450 1523
 FT CHAIN 451 1523
 FT CHAIN 452 1523
 FT CHAIN 453 1523
 FT CHAIN 454 1523
 FT CHAIN 455 1523
 FT CHAIN 456 1523
 FT CHAIN 457 1523
 FT CHAIN 458 1523
 FT CHAIN 459 1523
 FT CHAIN 460 1523
 FT CHAIN 461 1523
 FT CHAIN 462 1523
 FT CHAIN 463 1523
 FT CHAIN 464 1523
 FT CHAIN 465 1523
 FT CHAIN 466 1523
 FT CHAIN 467 1523
 FT CHAIN 468 1523
 FT CHAIN 469 1523
 FT CHAIN 470 1523
 FT CHAIN 471 1523
 FT CHAIN 472 1523
 FT CHAIN 473 1523
 FT CHAIN 474 1523
 FT CHAIN 475 1523
 FT CHAIN 476 1523
 FT CHAIN 477 1523
 FT CHAIN 478 1523
 FT CHAIN 479 1523
 FT CHAIN 480 1523
 FT CHAIN 481 1523
 FT CHAIN 482 1523
 FT CHAIN 483 1523
 FT CHAIN 484 1523
 FT CHAIN 485 1523
 FT CHAIN 486 1523
 FT CHAIN 487 1523
 FT CHAIN 488 1523
 FT CHAIN 489 1523
 FT CHAIN 490 1523
 FT CHAIN 491 1523
 FT CHAIN 492 1523
 FT CHAIN 493 1523
 FT CHAIN 494 1523
 FT CHAIN 495 1523
 FT CHAIN 496 1523
 FT CHAIN 497 1523
 FT CHAIN 498 1523
 FT CHAIN 499 1523
 FT CHAIN 500 1523
 FT CHAIN 501 1523
 FT CHAIN 502 1523
 FT CHAIN 503 1523
 FT CHAIN 504 1523
 FT CHAIN 505 1523
 FT CHAIN 506 1523
 FT CHAIN 507 1523
 FT CHAIN 508 1523
 FT CHAIN 509 1523
 FT CHAIN 510 1523
 FT CHAIN 511 1523
 FT CHAIN 512 1523
 FT CHAIN 513 1523
 FT CHAIN 514 1523
 FT CHAIN 515 1523
 FT CHAIN 516 1523
 FT CHAIN 517 1523
 FT CHAIN 518 1523
 FT CHAIN 519 1523
 FT CHAIN 520 1523
 FT CHAIN 521 1523
 FT CHAIN 522 1523
 FT CHAIN 523 1523
 FT CHAIN 524 1523
 FT CHAIN 525 1523
 FT CHAIN 526 1523
 FT CHAIN 527 1523
 FT CHAIN 528 1523
 FT CHAIN 529 1523
 FT CHAIN 530 1523
 FT CHAIN 531 1523
 FT CHAIN 532 1523
 FT CHAIN 533 1523
 FT CHAIN 534 1523
 FT CHAIN 535 1523
 FT CHAIN 536 1523
 FT CHAIN 537 1523
 FT CHAIN 538 1523
 FT CHAIN 539 1523
 FT CHAIN 540 1523
 FT CHAIN 541 1523
 FT CHAIN 542 1523
 FT CHAIN 543 1523
 FT CHAIN 544 1523
 FT CHAIN 545 1523
 FT CHAIN 546 1523
 FT CHAIN 547 1523
 FT CHAIN 548 1523
 FT CHAIN 549 1523
 FT CHAIN 550 1523
 FT CHAIN 551 1523
 FT CHAIN 552 1523
 FT CHAIN 553 1523
 FT CHAIN 554 1523
 FT CHAIN 555 1523
 FT CHAIN 556 1523
 FT CHAIN 557 1523
 FT CHAIN 558 1523
 FT CHAIN 559 1523
 FT CHAIN 560 1523
 FT CHAIN 561 1523
 FT CHAIN 562 1523
 FT CHAIN 563 1523
 FT CHAIN 564 1523
 FT CHAIN 565 1523
 FT CHAIN 566 1523
 FT CHAIN 567 1523
 FT CHAIN 568 1523
 FT CHAIN 569 1523
 FT CHAIN 570 1523
 FT CHAIN 571 1523
 FT CHAIN 572 1523
 FT CHAIN 573 1523
 FT CHAIN 574 1523
 FT CHAIN 575 1523
 FT CHAIN 576 1523
 FT CHAIN 577 1523
 FT CHAIN 578 1523
 FT CHAIN 579 1523
 FT CHAIN 580 1523
 FT CHAIN 581 1523
 FT CHAIN 582 1523
 FT CHAIN 583 1523
 FT CHAIN 584 1523
 FT CHAIN 585 1523
 FT CHAIN 586 1523
 FT CHAIN 587 1523
 FT CHAIN 588 1523
 FT CHAIN 589 1523
 FT CHAIN 590 1523
 FT CHAIN 591 1523
 FT CHAIN 592 1523
 FT CHAIN 593 1523
 FT CHAIN 594 1523
 FT CHAIN 595 1523
 FT CHAIN 596 1523
 FT CHAIN 597 1523
 FT CHAIN 598 1523
 FT CHAIN 599 1523
 FT CHAIN 600 1523
 FT CHAIN 601 1523
 FT CHAIN 602 1523
 FT CHAIN 603 1523
 FT CHAIN 604 1523
 FT CHAIN 605 1523
 FT CHAIN 606 1523
 FT CHAIN 607 1523
 FT CHAIN 608 1523
 FT CHAIN 609 1523
 FT CHAIN 610 1523
 FT CHAIN 611 1523
 FT CHAIN 612 1523
 FT CHAIN 613 1523
 FT CHAIN 614 1523
 FT CHAIN 615 1523
 FT CHAIN 616 1523
 FT CHAIN 617 1523
 FT CHAIN 618 1523
 FT CHAIN 619 1523
 FT CHAIN 620 1523
 FT CHAIN 621 1523
 FT CHAIN 622 1523
 FT CHAIN 623 1523
 FT CHAIN 624 1523
 FT CHAIN 625 1523
 FT CHAIN 626 1523
 FT CHAIN 627 1523
 FT CHAIN 628 1523
 FT CHAIN 629 1523
 FT CHAIN 630 1523
 FT CHAIN 631 1523
 FT CHAIN 632 1523
 FT CHAIN 633 1523
 FT CHAIN 634 1523
 FT CHAIN 635 1523
 FT CHAIN 636 1523
 FT CHAIN 637 1523
 FT CHAIN 638 1523
 FT CHAIN 639 1523
 FT CHAIN 640 1523
 FT CHAIN 641 1523
 FT CHAIN 642 1523
 FT CHAIN 643 1523
 FT CHAIN 644 1523
 FT CHAIN 645 1523
 FT CHAIN 646 1523
 FT CHAIN 647 1523
 FT CHAIN 648 1523
 FT CHAIN 649 1523
 FT CHAIN 650 1523
 FT CHAIN 651 1523
 FT CHAIN 652 1523
 FT CHAIN 653 1523
 FT CHAIN 654 1523
 FT CHAIN 655 1523
 FT CHAIN 656 1523
 FT CHAIN 657 1523
 FT CHAIN 658 1523
 FT CHAIN 659 1523
 FT CHAIN 660 1523
 FT CHAIN 661 1523
 FT CHAIN 662 1523
 FT CHAIN 663 1523
 FT CHAIN 664 1523
 FT CHAIN 665 1523
 FT CHAIN 666 1523
 FT CHAIN 667 1523
 FT CHAIN 668 1523
 FT CHAIN 669 1523
 FT CHAIN 670 1523
 FT CHAIN 671 1523
 FT CHAIN 672 1523


```

Db 738 KTEYKGVYDLTEGPEYFANGILTNLSYPSIIISHVNSPTLNREG--CCEYD--E 793
QY 472 APEYKHFRRCPFPFPKTVLERLLELRVAEMKKIPDPSEPRLLDEROKALKVLANA 531
Db 794 APQVGHFFCQDPQFPIBSLLDLDEROKVAKHKA-TVDPLEKKLLDYRGRALKILANS 852
QY 532 SYGFMGSGARWYCRECKAKATWANGRLHRTAI-NIARKLKLKVIY-----GDDP----- 580
Db 853 FYGIYGAAKARWYCKECKASVTANGROYIETTRKEIEKFGKVLVADSVTGDEVTIR 912
QY 581 ----- 580
Db 913 NGRIEFPVIEKLFERVDHVRGEKEYCVLGVEALTLNDRGLWKKVRYWRHHTDKRIY 972
QY 581 ----- 580
Db 973 RWFTNSWYLDVTEDEHSLIGYLTNSKVKPKRLERLVEVPEELGKVKSLITPNPPIA 1032
QY 581 ----- 580
1033 RTIQAAPLAVKLMELIGLVGDGKMGSGSNKATYVGLSCGLDKAEIERKVLNPLREASV 1092
QY 581 ----- 580
Db 1093 ISNYDKSKGQDVSLSKMLAGFMVKYFKDENGKKAIPSEFMNLPREYIEAFKGLPSAD 1152
QY 581 ----- 583
Db 1153 GTVSLRGIPEIRLTVSNRELSDAVRKLMLVGVNSLFTETKPNRYLEKESGTHSIHR 1212
QY 584 ----- 587
Db 1213 IKNGRPADRIIGFLDKRSTKLSNLSGHNKKQAYKTDPLVYPRKIEETLYGYVYDI 1272
QY 588 ----- 612
Db 1273 EVEGTHRFANGILVHNTDGFAPITPGADAFYKKAAREFLNINPKLPGILLEVESEFY 1332
QY 613 KRLPFTAKKRYAGLDEGRIDIVGEFAVRDGMCLAEVOTKVEIVLKTSEVNAKAVEY 672
Db 1333 RRGFFV-TKKCYAVVIDEKGKLTTRGLEIVRRDWSVAKETQARVLEALIRHGDVEAARI 1391
QY 673 VRKIVKELEBEGVPIEKLVIKTISKRLEETTEAPHVNAKMLSNAGYRSPDKIGYV 732
Db 1392 VKETEKLSKEVEPEKLVHEDQITREKDKYKATGPHALAKRLAARGIKVRPGTIVISYI 1451
QY 732 VRKIVKELEBEGVPIEKLVIKTISKRLEETTEAPHVNAKMLSNAGYRSPDKIGYV 732
1452 VLKSGRIGDRTIP-PDEFDPTRKRYDAEYIENOVLPAYERILKARGYKEDLRQKTR 1510
QY 791 QKTLFDPLAKSK 803
Db 1511 QVGLGAWLKMGKK 1523

```

```

RC STRAIN=KOD1;
RX MEDLINE=98027387; PubMed=9361436;
RA Takagi M., Nishioka M., Kakihara H., Kitabayashi M., Inoue H.,
RA Kawakami B., Oka M., Imanaka T.;
RT "Characterization of DNA polymerase from Pyrococcus sp. strain KOD1
RT and its application to PCR."
RL Appl. Environ. Microbiol. 63:4504-4510(1997).
RN [2].
RP CHARACTERIZATION OF INTENS.
RC STRAIN=KOD1;
RX MEDLINE=98416198; PubMed=9742242;
RA Nishioka M., Fujiwara S., Takagi M., Imanaka T.;
RT "Characterization of two intein homing endonucleases encoded in the
RT DNA polymerase gene of Pyrococcus kodakarensis strain KOD1."
RL Nucleic Acids Res. 26:4409-4412(1998).
CC -1- FUNCTION: INTEN ENCODED ENDONUCLEASES ARE THOUGHT TO MEDIATE
CC INTEN MOBILITY BY SITE-SPECIFIC RECOMBINATION INITIATED BY
CC ENDONUCLEASE CLEAVAGE AT THE "HOMING SITE" IN GENE THAT LACK THE
CC INTEN. PI-PKOI RECOGNIZES 5'-GATTAGATCCCTACC-3' AND PI-PKOII
CC RECOGNIZES 5'-CAGCTACTACGTTAC-3'. BOTH ARE THERMOSTABLE.
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + {DNA} (N).
CC -1- PTM: THIS PROTEIN UNDERGOES A PROTEIN SELF SPLICING THAT INVOLVES
CC A POST-TRANSLATIONAL EXCISION OF THE INTERVENING REGION (INTENIN)
CC FOLLOWED BY PEPTIDE LIGATION (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.
CC -1- SIMILARITY: IN THE INTENIN SECTION; BELONGS TO THE HOMING
CC ENDONUCLEASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, D29671; BAA06142.2; -.
DR HSSP; P56689; ITGO.
DR REBASE; 3792; PI-PKOI.
DR REBASE; 3793; PI-PKOII.
DR InterPro; IPR002064; DNA_pol_B.
DR InterPro; IPR003586; Hedgehog_hinc.
DR InterPro; IPR003587; Hedgehog_hinc.
DR InterPro; IPR002203; Inten.
DR InterPro; IPR004042; Inten_endonuc.
DR InterPro; IPR004578; Pol2.
DR Pfam; PF00136; DNA_pol_B_3.
DR Pfam; PF03104; DNA_pol_B_exo_1.
DR PRINTS; PR00379; INTENIN.
DR SMART; SM00305; HincC_2.
DR SMART; SM00306; HincC_2.
DR SMART; SM00486; POLBc_1.
DR TIGRFAMs; TIGR00592; pol2_2.
DR PROSITE; PS00116; DNA_POLYMERASE_B_1.
DR PROSITE; PS00818; INTENIN_C_TER_2.
DR PROSITE; PS00819; INTENIN_ENDONUCLEASE_2.
DR PROSITE; PS00817; INTENIN_N_TER_2.
KM Transferase; DNA-directed DNA polymerase; DNA replication;
KM DNA-binding; Hydrolase; Endonuclease; Autocatalytic cleavage;
KW Protein splicing; Inton homing.
FT CHAIN 1 406 DNA POLYMERASE, 1ST PART.
FT CHAIN 407 766 ENDONUCLEASE PI-PKOI.
FT CHAIN 767 851 DNA POLYMERASE, 2ND PART.
FT CHAIN 852 1388 ENDONUCLEASE PI-PKOI II.
FT CHAIN 1389 1671 DNA POLYMERASE, 3RD PART.
SQ SEQUENCE 1671 AA; 193490 MW; 4A17F1C8120BE455 CRC64;
Query Match 18.5%; Score 776; DB 1; Length 1671;
Best Local Similarity 17.6%; Pred. No. 1e-41;
Matches 304; Conservative 145; Mismatches 287; Indels 990; Gaps 24;
QY 8 VLDSYEVVKEQVVIWGIAENGERVVLIDRSFRFYFALLAPGADPKQV-----AQRIR 63

```

```

Db      2  ILDTDYITEDGKPVIRIFK--KENGEFKIEVDRTFEFFYFALLKODSALIEVKKITTERRIG 60
Qy      64  ALSRPSPILIGVEDDKKCYFGRPRVLRJRTVLPEAVREYELVANDGVEDLVEDITF 123
Db      61  TVTVYGR-----VEKQKQKFLGRFVEVWKLFTHPOVRAIRDKIREHNAVIDIYEDITF 116
Qy      124  ARMYLIDHDLFFFTWYRVEAPLENNKGFVNDKTVLVSRRPELYGALAPTLPLDRL 183
Db      117  AKKYLIDKGLV-----PMEGD-----EELKML 138
Qy      184  APDIEVYSKQSPRPERDPVIVIAKTDGDEVL-----FLAEGDKDRKPIREFVEY 235
Db      139  AFDIETLYHEGEEFAL--GPILMI SYADEGARVITWKNVDLPYVDVSTEREMIKKFLV 197
Qy      236  VKRYDPDIIIVGNNNHPPMPYLLRRARILGIKLDVTRVGAAPTTSVGH---VSYPGL 292
Db      198  VKEKDPDVLITNGDNDFALYKRRCEKLGINFALGRD--GSBPKIQHWGDRFAVEYKGR 256
Qy      293  NVDLYDAEEMPEIKIKSLSEVAYLGVMKKSEFVIMMWEIPYMDPPKRPILLQYAR 352
Db      257  HFDLYEVIRRTINLPYTLLEAVYEAVGQPK--EKVYAE--BITTAMETGENLERVARSEM 313
Qy      353  DDPVATYGLAEKILPFAIQLSYVTGLPDQVGMASVGRLEMYLIRAPFKKELVPRVE 412
Db      314  EDKATYELGKKEFLPMEQJLSRLIGOSLMVSRSSITGNLVEWFLIRKQYERNEIAPKED 373
Qy      413  RPE-----ETYGALVLEPLAGVHENAVLDF----- 439
Db      374  EKELARRQSTIEGGYVKEPEKGMENIVYDFRCHPADTKVVVKGKGINISEVQGDYV 433
Qy      440  ----- 439
Db      434  LGIDGQVRKVMYDYKGLVNINGLKCTPNHKL.PVYTKOERQTRIRDSIANS.FLTKEY 493
Qy      440  ----- 439
Db      494  KGKILTPLEIGRATSENIPBEVYLKGLAGIILAEGTLRKDVEFDSHKKRRISH 553
Qy      440  ----- 439
Db      554  QYRVEITIGDEEPRDRITYIFERLFGITPSISEKKGNATVLKAKKQVYLKVEIMD 613
Qy      440  ----- 439
Db      614  NIESLHAPSVLRGFEEDGSVNRVRSIVATGCTKNEKIKLVSKLSQGLGHQYTYQ 673
Qy      440  ----- 439
Db      674  YQENGKDRRYLLETGKGLILFQTLIGISERKALLNKAISOREMNLENNRYRUS 733
Qy      440  ----- 467
Db      734  EFNVSTEYEGKVNDLLEGTIPPYFANGILTHNSLYPSITTHNVSDPTLNR----- 786
Qy      468  GCNE--AEEVGRFRRCPPGFVTLERLLBELRKRAEMKKYPPSPFYRLIDRQKA 524
Db      787  GCKEYDVAPOVGRHPCDPBGFIPSLIGLLEBKOKIKKKKA--TIDPIERKLLDYROBA 845
Qy      525  LKVL----- 529
Db      846  IKILANSILPEEMLPVLEEGEVHFRIGELIDRMENAGKVGEGTEVLVSGLEVPS 905
Qy      530  ----- 529
Db      906  FNRRTNKALKRVKALRHODSGKVYTIKSGRRIKITSQHSLEPSVANGELVEYVGDLE 965
Qy      530  ----- 529
Db      966  KPEDLVAVPRLELPERNHVLNVLLELLGTPEEETLDIVMTIPVKGKKNFKKMLRTLNR 1025
Qy      530  ----- 529

```

```

Db      1026  IFGEERPRPTARVIRHLEDLGYYRLKKIYCEVLDMDSLAKYRRLYALVENRYNNGKR 1085
Qy      530  ----- 529
Db      1086  EYLVENSTIRDAVGMPLKELKEMKIGTLNGFRMRKLIEVDESLAKLLGYVSEGYARQ 1145
Qy      530  ----- 529
Db      1146  RNPANGMSYKLVINEDEVLDMERLASFGKVRGRGNYVEIPKKI GYLLFENNCGVL 1205
Qy      530  ----- 529
Db      1206  AENKRIPEFVTPKGVLAFLLEGYFIGDDVHPNKRRLSTKSELLANQVLILNSGV 1265
Qy      530  ----- 529
Db      1266  SAVKLHDSGVYRYINEELPFVYLDKKNAVYSHVIPREVLSVFGKVFQKVPQTR 1325
Qy      530  ----- 529
Db      1326  KMWEDRLDPEKQRLSLIEGDVYLDREVSVDEVDGYTYDLSVEDNENFLVGRGLY 1385
Qy      530  --NASYGVNMGSGARVYREGAKAYTANGRHLIRAI--NIARKLGLVYIGDDSLFPTY 586
Db      1386  AHSYIGYVAPAMWYCKEASVTMGREYITMTIKEIERKGFVITSDDGFFAT 1445
Qy      587  ---DPEVE---NFIKTIKEEL--GEIKLEKVRQLPFTBAKKRYAGLLEDGRIDIVG 637
Db      1446  PGADAETVKKAMEFLKYINKLPGALELEYEGFPYKRGFPV--TKKRYAVDEGKLTTRG 1504
Qy      638  FEAVRGDMCELAKEVQTKVEIYVLTSEVNAKAVEYRKIVLEBEGVPIEKLVIKTL 697
Db      1505  LEIVRDMSEIAKEQAVLEALLKQGDVEKVRIVKEVTEKLSKYEVPPEKLVIEHQIT 1564
Qy      698  KRLEEYTPAPHVNAKMLSGAGRVSPQDKIGYIVVAGGRISQAPAPFYMKDPS--Q 755
Db      1565  RDLKDYKATGPVAVAKKLARGVKIRPGTVISYIVLKGSRIGDRAP--FDFPDPTK 1623
Qy      756  IDVTYVHDQIIPALRILIGYFGITEKCLKASATGQKTLPPFLAK 801
Db      1624  YDAEYIENQVPAVERILIRAFGRKEDLRKYQTKQVGLSMLMPK 1669

RESULT 15
PROD SCHPO STANDARD; PRT; 1086 AA.
AC F30316; Q10016; Q9USU0;
DT 01-APR-1993 (Rel. 25, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DNA polymerase delta catalytic subunit (EC 2.7.7.7) (DNA polymerase
DE 11).
GN POL3 OR POLD OR SPBC36.04.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92071954; PubMed=1960723;
RA Plagnade G., Bouvier D., de Recondo A.M., Baldacci G.,
RT "Characterization of the Pol3 gene product from Schizosaccharomycetes
RT pombe indicates inter-species conservation of the catalytic subunit
RT of DNA polymerase delta."
RL J. Mol. Biol. 222:209-218(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93184400; PubMed=8443413;
RA Pak H., Francesconi S., Wang T.S.F.,
RT "Cell cycle expression of two replicative DNA polymerases alpha and
RT delta from Schizosaccharomycetes pombe."
RL Mol. Biol. Cell 4:145-157(1993).

```

[3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgourou J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holtvyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jørgels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Maier P., Moles S., Mungall K., Murphy L., Nisbett D., Odeli C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Stevens K.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Sharp S.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Voiclaekart G., Aert R., Robben J., Grynompres B.,
 RA Weltjens I., Vantretels R., Rieger M., Schefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lettrach H., Reinhardt R., Pohl T.M.,
 RA Egger P., Zimmermann W., Medler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu B., Dreano S., Gloux S., Leleau V., Motier S.,
 RA Gaibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Useary D., Barrett B.G., Nurse P.;
 RL "The genome sequence of *Schizosaccharomyces pombe*.";
 CC Nature 415:871-880(2002)
 CC -1- FUNCTION: THIS POLYMERASE POSSESSES TWO ENZYMAIC ACTIVITIES:
 CC DNA SYNTHESIS (POLYMERASE) AND AN EXONUCLEOLYTIC ACTIVITY THAT
 CC Degrades SINGLE STRANDED DNA IN THE 3' TO 5' DIRECTION.
 CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
 CC + {DNA} (N).
 CC -1- SUBUNIT: HETEROETETRAMER THAT CONSIST OF THE POL3, CDC1, CDC27 AND
 CC CCM1 SUBUNITS. THE POL3 SUBUNIT CONTAINS THE POLYMERASE ACTIVE
 CC SITE AND MOST LIKELY THE ACTIVE SITE FOR THE 3'-5' EXONUCLEASE
 CC ACTIVITY.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- MISCELLANEOUS: IN EUKARYOTES THERE ARE FIVE DNA POLYMERASES:
 CC ALPHA, BETA, GAMMA, DELTA, AND Epsilon WHICH ARE RESPONSIBLE FOR
 CC DIFFERENT REACTIONS OF DNA SYNTHESIS.
 CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation --
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.gsf).
 CC -----
 CC CC
 CC EMBL; X52978; CAA41968.1; --
 CC EMBL; U07734; AAA35303.1; --
 CC EMBL; AL121815; CAB58156.1; --
 CC PIR; S19661; S19661.
 CC InterPro; IPR002064; DNA_pol_B.
 CC InterPro; IPR004578; POL2.
 CC Pfam; PF00136; DNA_pol_B; 1.
 CC Pfam; PF03104; DNA_pol_B_exo; 1.
 CC PRINTS; PR00106; DNAPOB.
 CC SMART; SM00486; POLB.
 CC TIGRFAMs; TIGR00592; pol2; 1.
 CC PROSITE; PS00116; DNA_POLYMERASE_B; 1.
 CC DR Transferrase; DNA-directed DNA polymerase; DNA replication;
 CC KM DNA-binding; Hydrolase; Exonuclease; Zinc-finger; Nuclear protein.
 CC FT ZN_FING 993 1011 C4-TYPE (POTENTIAL).
 CC FT ZN_FING 1040 1058 Q -> E (IN REF. 1).
 CC FT CONFLICT 102 102 O -> E (IN REF. 1).
 CC FT CONFLICT 419 419 T -> S (IN REF. 1).
 CC FT CONFLICT 545 545 R -> C (IN REF. 1 AND 2).
 CC FT CONFLICT 777 784 KLEFEKRY -> NMSFST (IN REF. 1).
 CC FT CONFLICT

```

FT  CONFLICT  866      866      L - H (IN REF. 1)
SQ  SEQUENCE  1086 AA; 123568 MM; 99952841322003CA CRC64;

Query Match      18.4%; Score 771.5; DB 1; Length 1086;
Beet Local Similarity 27.3%; Pred. No. 1.2e-41;
Matches 246; Conservative 160; Mismatches 319; Indels 177; Gaps 30;

QY  3  EVFVPLVDSSEVVGKEQVIVWIGIAENGERVVLIDNSFRPYFALLAPGADPKVOAORI 62
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB  97  DIFQOIDSSEFTEGSVPSIRLFGVTDGNSILVHVGFPLPYFVKAQVGRPREMLERFT 156
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY  63  RAL-----SRPSPIIGVDDBRKVFGRPRVLRIRTVLPE 98
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB  157  QDLATCNGVYIDHCITIMKENLYGFOGNEKSPFIKI-----FTTNRILSRANVPER 210
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY  99  AVREYRELVKQVGDVEDVLEADIRFAMRYLLIDHDLFFPTWRYAEAPLENKMGFEV---- 154
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB  211  GEFNEELFPVGVGV-TTFESNTGYLARFMIDCVGVGNMVIHLPRASKQGFYQNRVSNQC 269
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY  155  -----DKYLVKSRPEPLVGEALAPTKLPDLRLAFDIEVYSKQG-SRPERDPYIV 205
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB  270  IEAWINYKDLISL-----PAGGW---SKMAPLRIMSFIDECAGRKGVFPDPIDPYIQ 320
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY  206  IA-VTDDGDEVLFIA-----EKKDRKPIREVEVVKXYPDIITYG 247
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB  321  IASITQYGDSTPFVRNVFCVDTCSQIVGTQVYEBQNAEMLSMSKFRVDVDPVLLGY 380
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY  248  NNNHFDPMYLLRRAILGI-----KLDVTRVCAEPTTSVHGH-----VSVGRSLN 293
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB  361  NICNDPIYLLDRAKSLAIHNPPLLRGHRHFFSVAKETTFSSKAYGTRESEKTTISIPGLQ 440
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY  294  VDLVYDAEEMPEIKISLEEV-AEYLGVWKKSERVIINWMEIPTYDW---DPKRRPLLQ 349
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB  441  LDMLOVQORDFKLSYSINAVCSOPLGROKED---VHYSITTDLQNGTADSRRR--LAI 494
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY  350  YARDVVRATYGLAEILTEFA--IQLSYVTGLPQGVGMSGFRLEWLLIAPFCKELV 407
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB  495  YCLKNAVYLPQRIMDLTMCFVNTTEARATVGPVFNFLARGOQIKVLSQLFRKALQHDLVV 554
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY  408  PN-RVERDEETRYGALVLEPLRGVHEN-IAVLDEFSMTPNIM----- 447
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB  555  PRIRNGTDEQEGATVIEPIKGYIDTIALDLDSLSLYPSIMQAHNLCTYLLDSNTAEL 614
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY  448  ----IKYVGP--DTLVPRGEKCGEGCWAEABEVKHRFRRCPPGFEKTVLERLLELRK 499
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB  615  LKLKQDVVYSVTPNGDYPFVKPHVR-----KGLLPILLADILNARK 654
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY  500  RVRAEMKRYPPDSPEYRLLDEROKALKYLANASGYGMSGARMTCRECAVATVAMGHL 559
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB  655  KAKADLKR-ETDPFKKAVLDRQALAKTSANSVGFTGATNGRLPCLAISSSVISYGRQM 713
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY  560  IRTAINIRKL-----GLKVIYGDTSLEFVYDEKVENFIKIIKEELGF----- 604
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB  714  IEXTQVVEKRYRIENGSHDAVVIYGDTSVMKFGYKTLPEAKKIGEBEANYSDQFP 773
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY  605  ---EIKLEKVKYRLFTEAKKRYAGLL-----EDGRIDIVFEFAVRGDWCELAKEVQTV 656
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB  774  NPRIKLEFQVFP-YLISKRRVYGLFMTRTDTYDKMDSKSIETVRBNCLVSVVIDTA 832
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY  657  VEIVLTSEVNAKAYEVYKVIKYLEEGVPIEKLVIWKTLSKRLSEYTTAPHYVAARM 716
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB  833  LRKMLIDDDVGAQOLFTEKKVISDLQNKICIMSOAVITKALSK--TDVAAKMAHVELAERM 890
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY  717  --LSAGVNSQGDKIGVYIVG--GGRISORAM-PYFVVKQPSQIDPYVYVDHQIIPAL 771
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB  891  KRKDGSAIPAIGDRAVAVIILKGAQGDQFYKRSDEPIYLENNIPIDAKYILENLSKPL 950
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY  772  RI 773
    : :
DB  951  RI 952
    : :

```

Wed Apr 23 08:09:18 2003

Job time : 28 secs

us-10-034-849-2.rsp

Page 16

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 07:28:09 ; Search time 41 Seconds

(without alignments)
4035.508 Million cell updates/sec

Title: US-10-034-849-2

Perfect score: 4188
Sequence: 1 MTEVFTVLDSSYEYVVKEP.....LKASATGQKTLPPFLAKSK 803

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database:

SPTREMBL_21:*

- 1: sp archaea:*
- 2: sp bacteria:*
- 3: sp fungi:*
- 4: sp human:*
- 5: sp invertebrate:*
- 6: sp mammal:*
- 7: sp mhc:*
- 8: sp organelle:*
- 9: sp phage:*
- 10: sp plant:*
- 11: sp rodent:*
- 12: sp virus:*
- 13: sp vertebrate:*
- 14: sp unclassified:*
- 15: sp virus:*
- 16: sp bacteriophage:*
- 17: sp archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3033.5	72.4	803	1	Q59691
2	2248	53.7	781	17	Q96VY1
3	2246	53.6	781	1	Q96VY1
4	2156	51.5	785	17	Q8ZVY2
5	2129.5	50.8	785	1	Q96VY2
6	1268	30.2	775	1	Q96VY4
7	1264	30.2	775	1	Q96VY4
8	1220	29.1	783	17	Q8TWJ5
9	1208	28.8	759	1	Q96VY1
10	852.5	20.4	937	17	Q8TSB3
11	760.5	18.2	1107	4	Q96VY8
12	743.5	17.8	1105	11	Q91VT0
13	737.5	17.6	1104	3	Q8X0V7
14	732.5	17.5	974	5	Q8SQP5
15	718.5	17.2	901	17	Q9HRV9
16	655.5	15.7	855	17	Q8ZVY9

17	649.5	15.5	914	1	Q59690	Q59690 pyrodicticum
18	648	15.5	876	17	Q971C7	Q971C7 sulfobolus
19	646.5	15.4	796	17	Q97AH0	Q97AH0 thermoplasm
20	622	14.9	800	17	Q97AH3	Q97AH3 thermoplasm
21	575.5	13.7	1458	13	Q9DE46	Q9DE46 xenopus lae
22	569	13.6	787	16	Q9KSP4	Q9KSP4 vibrio chol
23	564.5	13.5	844	1	Q31096	Q31096 cenarchaeum
24	560.5	13.4	1016	12	Q993K6	Q993K6 callitrichi
25	552	13.2	845	1	Q74046	Q74046 cenarchaeum
26	540	12.9	1492	10	Q9FHA3	Q9FHA3 arabidopsis
27	539	12.9	2992	4	Q9N25	Q9N25 homo sapien
28	536	12.8	1015	12	Q8UZD7	Q8UZD7 cercopithec
29	527.5	12.6	787	16	Q91ZL1	Q91ZL1 pseudomonas
30	524	12.5	1013	12	Q91LX9	Q91LX9 retroperito
31	521.5	12.5	1013	12	Q41208	Q41208 retroperito
32	515	12.3	956	5	Q8SRW2	Q8SRW2 pseudorabito
33	511	12.2	956	5	Q8SRW2	Q8SRW2 encephalito
34	511	12.2	1026	12	Q36363	Q36363 alceaphine
35	510	12.2	1004	12	Q96Z8	Q96Z8 porcine lym
36	507.5	12.1	1443	5	Q9NKH1	Q9NKH1 caenorhabdi
37	506.5	12.1	1012	12	Q40915	Q40915 kaposi's sa
38	506.5	12.1	1012	12	Q40910	Q40910 kaposi's sa
39	506.5	12.1	1472	5	Q9VD90	Q9VD90 drosophila
40	506	12.1	774	2	Q9F175	Q9F175 pseudomonas
41	503	12.0	1041	12	Q9PKT8	Q9PKT8 elephant he
42	501	12.0	1005	12	Q99D20	Q99D20 bovine hept
43	501	12.0	1009	12	Q9YTQ4	Q9YTQ4 atelina her
44	499.5	11.9	789	16	Q8ZIH9	Q8ZIH9 yersinia pe
45	497	11.9	1415	5	Q77034	Q77034 drosophila

ALIGNMENTS

RESULT 1

ID Q59691 PRELIMINARY; PRT; 803 AA.

AC Q59691;

DT 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE DNA polymerase (EC 2.7.7.7).

GN POLB

OS Pyrodicticum occultum.

OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcaceae;

OC Pyrodictiaceae; Pyrodicticum.

OX NCBI_Taxid=2309;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=DSM2709;

RX MEDLINE=95238290; PubMed=721707;

RA Demori T., Ishino Y., Doi H., Kato I.;

RT "The hyperthermophilic archaeon Pyrodicticum occultum has two alpha-like DNA polymerases";

RT J. Bacteriol. 177:2164-2177(1995).

RL J. Bacteriol. 177:2164-2177(1995).

CC - - CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE = N DIPHOSPHATE + {DNA} (N).

CC - - SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.

CC EMBL; D38574; BAA07580.1; -.

DR HSSP; P56689; ITGO.

DR InterPro; IPR002064; DNA_pol_B.

DR InterPro; IPR004578; Pol2.

DR Pfam; PF00136; DNA_pol_B_1.

DR Pfam; PF03104; DNA_pol_B_exo_1.

DR PRINTS; PR00106; DNAPOLB.

DR SMART; SM00486; POLBc; 1.

DR TIGRPFAM; TIGR00592; pol2; 1.

DR PROSITE; PS00116; DNA_POLYMERASE_B_1.

DR DNA replication; DNA-binding; DNA-directed DNA polymerase;

KW Nucleotidyltransferase; Transferase.

SW SEQUENCE 803 AA; 92656 MW; E9C092F26A8D23FA CRC64;

Query Match 72.4%; Score 3033.5; DB 1; Length 803;

Best Local Similarity 70.9%; Pred. No. 6,86-193;
Matches 571; Conservative 106; Mismatches 121; Indels 7; Gaps 5;

```

QY 1 MTEVY-FTVLDSSEYVNGKEPOVLIINGIAENGERVYLIDRSFRPFYALLAPGADP--KO 57
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MTEIEFVLDSSEYVNGKEPOVLIINGIAENGERVYLIDRSFRPFYALLAPGADP--KO 60
QY 58 VAGRTAASRSPKSPIGVEDDKKRCGRPRVRLRTVLPAAVREYELVKNVGVEDVL 117
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 IASIRLSVKSPIIDAKPDKKTRFPAKAKITIMLPSSVHTYBAVKLEGGVDSL 120
QY 118 EADIRFANRYLIDHDLFFFTYRVEAEPLNNKGRPRVDKYLVSREPLYGALAPTKL 177
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 EADIRFANRYLIDHDLFFFTYRVEAEPLNNKGRPRVDKYLVSREPLYGALAPTKL 178
QY 178 PDLRIAPDIEVYSKQSPREPRDPVIVIAVKTDDDEVLFIAEGKDRKPIREFEYVK 237
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 179 PPKRLVAPDIEVYSKQSPREPRDPVIVIAVKTDDDEVLFIAEGKDRKPIREFEYVK 238
QY 238 RYDPDIIYGVNNNHDPYLLRLARIGILDTYTRVGAEPFTSVHGVSVGPRLAVDLY 297
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 239 ACPDPIYGVNNNHDPYLLRLARIGILDTYTRVGAEPFTSVHGVSVGPRLAVDLY 298
QY 298 DVAEEMPEIKRSLEEVAYELGVWKSSEVLIINMWEIDYMDDEKRPILLOYARDVNA 357
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 299 DVAEEMPEIKRSLEEVAYELGVWKSSEVLIINMWEIDYMDDEKRPILLOYARDVNA 358
QY 358 TYGLAEKILPRLIQLSVYTGVLPLDVGANSVGRLEWYLIRAAKMKELVPRNVERPEET 417
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 359 TYGLAEKILPRLIQLSVYTGVLPLDVGANSVGRLEWYLIRAAKMKELVPRNVERPEET 418
QY 418 YKGAIVLEPLKGVHENIAVLDPSMTPTNIMIKYNGPPTLVPRSEKSGEC--GCWEAPEVK 476
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 419 YKGAIVLEPLKGVHENIAVLDPSMTPTNIMIKYNGPPTLVPRSEKSGEC--GCWEAPEVK 478
QY 477 HRRRCPPGFEKTVLERLLERLRKRVAKKYPSPDSPEYRLIDEROALKVLNANASYGM 536
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 479 HRRRCPPGFEKTVLERLLERLRKRVAKKYPSPDSPEYRLIDEROALKVLNANASYGM 538
QY 537 GWSGARWYCRECAKAVTAMGSHLIRPAINIARLGLKAYIGDTSLSFTYDPPKVNPFK 596
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 539 GWSGARWYCRECAKAVTAMGSHLIRPAINIARLGLKAYIGDTSLSFTYDPPKVNPFK 598
QY 597 IIEKELGEFIEKLEKYKRLPFTBAKKRYAGLJEDGRIDIVGEAVRGMDELAKVOTRV 656
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 599 IIEKELGEFIEKLEKYKRLPFTBAKKRYAGLJEDGRIDIVGEAVRGMDELAKVOTRV 658
QY 657 VEVLKTSRVNKAVEYRKVYKLEBGRPIELVYKTLISKRLSEETTEAPRVVAAKRM 716
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 659 VEVLKTSRVNKAVEYRKVYKLEBGRPIELVYKTLISKRLSEETTEAPRVVAAKRM 718
QY 717 LSAGRVSPGDKIGVYKGGGRISQAMPYEVNDPSQIDVTYVYDHQIIPALRIIGY 776
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 719 LSAGRVSPGDKIGVYKGGGRISQAMPYEVNDPSQIDVTYVYDHQIIPALRIIGY 777
QY 777 FGITTEKCLKASATGKTLFDFLAKK 801
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 778 FGITTEKCLKASATGKTLFDFLAKK 802
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

RESULT 2

Q96YV1

ID Q96YV1

PRELIMINARY;

PRT;

781 AA.

AC Q96YV1

DT 01-DEC-2001

(Tremblere1, 19, Last sequence update)

DT 01-DEC-2001

(Tremblere1, 19, Last sequence update)

DE Putative DNA-directed DNA polymerase II.

GN ST2076.

OS Sulfolobus tokodaii.

OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;

OC Sulfolobus.

OX NCBI_taxid=111955;

```

RN [1]
RP SEQUENCE FROM N.A.
RC SPTAIN=JOM 10545 / 7;
RX PubMed=11572479;
RA Kawanabe Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Anka H., Kobayashi H., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudo Y., Yamazaki J., Koshida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermophilic
RT Crenarchaeon, Sulfolobus tokodaii strain7."
RL DNA Res. 8:123-140(2001).
DR EMBL; AP000988; BAB67175.1; -.
DR InterPro; IPR002064; DNA_pol_B.
DR Pfam; PF00136; DNA_pol_B; 1.
DR TIGRFam; TIGR00592; pol2; 1.
DR PROSITE; PS00116; DNA polymerase; Hypothetical protein; Complete proteome.
KW DNA-directed DNA polymerase; Hypothetical protein; Complete proteome.
SQ
SEQUENCE 781 AA; 90447 MW; 25E61383DDF903EF CRC64;

Query Match 53.7%; Score 2248; DB 17; Length 781;
Best Local Similarity 54.0%; Pred. No. 9,4e-141;
Matches 430; Conservative 143; Mismatches 203; Indels 20; Gaps 10;

QY 6 FTVLDSSYEVNGKEPOVLIINGIAENGERVYLIDRSFRPFYALLAPGADPQVAAQRIRAL 65
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 5 FTVLDSSYEVNGKEPOVLIINGIAENGERVYLIDRSFRPFYALLAPGADPQVAAQRIRAL 64
QY 66 SRKSPPIIGVEDDKKRYGRRRLRLRTVLPAAVREYELVKNVGVEDVLADIRFAM 125
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 65 SRKSPPIIGVEDDKKRYGRRRLRLRTVLPAAVREYELVKNVGVEDVLADIRFAM 124
QY 126 RYLDHDLFFFTYRVEAEPLNNKGRPRVDKYLVSREPLYGALAPTKLPLRLIAP 185
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 125 RYLDHDLFFFTYRVEAEPLNNKGRPRVDKYLVSREPLYGALAPTKLPLRLIAP 177
QY 186 DIEVYSKQSPREPRDPVIVIAVKTDDDEVLFIAEGKDRKPIREFEYVKYVDDPIIV 245
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 186 DIEVYSKQSPREPRDPVIVIAVKTDDDEVLFIAEGKDRKPIREFEYVKYVDDPIIV 245
QY 246 GYNNHDPYLLRLARIGILDTYTRVGAEPFTSVHGVSVGPRLAVDLYDVAEMPE 305
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 246 GYNNHDPYLLRLARIGILDTYTRVGAEPFTSVHGVSVGPRLAVDLYDVAEMPE 305
QY 297 VKKTELENIADVLGVLPEKRTIVEMWDIPKWDDEKRRDILKKNLDDAKSAVLLGEVF 356
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 297 VKKTELENIADVLGVLPEKRTIVEMWDIPKWDDEKRRDILKKNLDDAKSAVLLGEVF 356
QY 306 IKRSLEVAEYLVGWSKSEVLIINMWEIDYMDDEKRPILLOYARDVATYGLAEKI 365
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 306 IKRSLEVAEYLVGWSKSEVLIINMWEIDYMDDEKRPILLOYARDVATYGLAEKI 365
QY 366 LPFALQLSVYTGVLPLDVGANSVGRLEWYLIRAAKMKELVPRNVERPEETRGAILVE 425
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 366 LPFALQLSVYTGVLPLDVGANSVGRLEWYLIRAAKMKELVPRNVERPEETRGAILVE 425
QY 357 IPGIEIEITRISGLPLDQLSMASVGRVEMLRKAEAKYVELLPNNEEVEYESGGLVTS 416
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 357 IPGIEIEITRISGLPLDQLSMASVGRVEMLRKAEAKYVELLPNNEEVEYESGGLVTS 416
QY 426 PLGCVHENIAVLDPSMTPTNIMIKYNGPPTLVPRSEKSGEC--GCWEAPEVKRPRCP 484
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 426 PLGCVHENIAVLDPSMTPTNIMIKYNGPPTLVPRSEKSGEC--GCWEAPEVKRPRCP 484
QY 417 PLGCIHEDVYVLDPSMTPTNIMIKYNGPPTLVK-----GCEENQWISP--AGKHFKKDP 470
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 417 PLGCIHEDVYVLDPSMTPTNIMIKYNGPPTLVK-----GCEENQWISP--AGKHFKKDP 470
QY 485 GFEKTVLERLLERLRKRVAKKYPSPDSPEYRLIDEROALKVLNANASYGMGSGARV 544
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 485 GFEKTVLERLLERLRKRVAKKYPSPDSPEYRLIDEROALKVLNANASYGMGSGARV 544
QY 471 GLKYVLEKLIORKEVKGKMER--TMDYKRVLDKRGKALVANAFAFYGMGLGARNY 529
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 471 GLKYVLEKLIORKEVKGKMER--TMDYKRVLDKRGKALVANAFAFYGMGLGARNY 529
QY 545 CRECAKAVTAMGSHLIRPAINIARLGLKAYIGDTSLSFTYDPPKVNPFK 604
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 545 CRECAKAVTAMGSHLIRPAINIARLGLKAYIGDTSLSFTYDPPKVNPFK 604
QY 530 SKGAEAVTAMGSHLIRPAINIARLGLKAYIGDTSLSFTYDPPKVNPFK 588
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 530 SKGAEAVTAMGSHLIRPAINIARLGLKAYIGDTSLSFTYDPPKVNPFK 588
QY 605 EIKLEEVYKRLPFTBAKKRYAGLJEDGRIDIVGEAVRGMDELAKVOTRVVEIYKTS 664
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 605 EIKLEEVYKRLPFTBAKKRYAGLJEDGRIDIVGEAVRGMDELAKVOTRVVEIYKTS 664
QY 589 EIKLIDKTVKRVFTENKRRYAGLJEDGRIDIVGEAVRGMDELAKVOTRVVEIYKTS 648
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 589 EIKLIDKTVKRVFTENKRRYAGLJEDGRIDIVGEAVRGMDELAKVOTRVVEIYKTS 648
QY 665 EVNKAVEYRKVYKLEBGRPIELVYKTLISKRLSEETTEAPRVVAAKRMISAGYRS 724
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 665 EVNKAVEYRKVYKLEBGRPIELVYKTLISKRLSEETTEAPRVVAAKRMISAGYRS 724

```

Db 649 KVEDAIKVVKTVIPDLRRYNEFIEDLIIMKTIDKNLDEYVATAPHVAAKKAAGYLV 708

Qy 725 PGDKIGVIVYKGGRIISQRAPIYFMVNDPSQIDPTVYVDHDIIPALRIIGYFITEKKL 784

Db 709 KGVIGIVYKGGRIISQRAPIYFMVNDPSQIDPTVYVDHDIIPALRIIGYFITEKKL 768

Qy 785 KASATGOKTLPDFLAK 800

Db 769 K---TGVDILSPFK 781

RESULT 3

Q9P9N1 PRELIMINARY; PRT; 781 AA.

ID Q9P9N1

AC 01-OCT-2000 (Tremblrel. 15, Created)

DT 01-OCT-2000 (Tremblrel. 15, last sequence update)

DT 01-JUN-2002 (Tremblrel. 21, last annotation update)

DE DNA polymerase B3.

NC Sulfuriphraera ohwakensis.

OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;

OK Sulfuriphraera.

OK NCBI_TaxID=69656;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=TA-1;

RX MEDLINE=20450680; PubMed=10997874;

RA Imai T., Kurohara N., Itoh Y.H., Kimura N., Horinchi T.;

RT "Sequence analysis of three family B DNA polymerases from the

RT thermophilic crenarchaeon Sulfuriphraera ohwakensis."

RL DNA Res. 7:243-251(2000).

CC -1- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE = N DIPHOSPHATE + [DNA] (N).

CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.

DR EMBL; AB032376; BAA93703.1; -.

DR HSSP; P56689; 1TGO.

DR InterPro; IPR002064; DNA_pol_B.

DR InterPro; IPR004578; Pol2.

DR Pfam; PF00136; DNA_pol_B; 1.

DR Pfam; PF03104; DNA_pol_B_exo; 1.

DR PRINTS; PR00106; DNAPOLB.

DR SMART; SM00486; POLBc; 1.

DR TIGRFAMs; TIGR00592; pol2; 1.

DR PROSITE; PS00116; DNA_POLYMERASE_B; 1.

KM DNA replication; DNA-Binding; DNA-directed DNA polymerase.

SO SEQUENCE 781 AA; 90384 MW; 382570EAB30C372C CRC64;

Query Match 53.6%; Score 2246; DB 1; Length 781;

Best Local Similarity 54.3%; Pred. No. 1,38-140;

Matches 432; Conservative 141; Mismatches 203; Indels 20; Gaps 10;

Qy 6 FTUDDSTYEVGKEPOVIWIGIAENGSEFVILIDRSFRYPALALPGADPKQVARIATL 65

Db 5 PFILDFSYDVENKPRVYIWIWIDKGNVVLLEKKFRYPALVDDSYNIDEIRKILKL 64

Qy 66 SRKPSPIGVDDKRYKGRBRVYLRITVYBEAVREYRELVKNDVGVDELADIRPAM 125

Db 65 SKPSPISISIVEEKTFGSPVKALKITVYPAYVRVDRVAKIKGVKSLVLEADIRYM 124

Qy 126 RYLIDHDLFPFTWYRVEAPLENKGFPRVDKYLVYKSRPELYGALAPTLPLRIALF 185

Db 125 RYSDINLKPFYWEAEVEIKEN-NFRVKKVYELK-KINTLYD-----KIPELKYLAF 177

Qy 186 DIETVSKSGSRPERDPYIVAVKTDDGDEVLFIAEGDKRPPREFEYVKRYDPDIIV 245

Db 178 DIEVYNNKSGSNPRDPYIIGVWTKESGK-QFLADKDDLRARIREFTNFVQVYDPDIIV 236

Qy 246 GYNNHHPDMPYLRRARLIGIKLQVTRVGAEPSTVGHVSVGRLNVDLYDAEEMPE 305

Db 237 GYNNHHPDMPYLRRARLIGIKLQVTRVGAEPSTVGHVSVGRLNVDLYDAEEMPE 296

Qy 306 IKISLEEVAVYLGVMKSESVIINWMEIPDYWDPKRRPILLOYARDVATYGLAEKI 365

Db 297 VKVKTLENIADYIGLVPEKERTIVEMWDIPKWNDEKKRIDILKNLDDAKSAVILGEVF 356

Qy 366 LPPAIQLSYVTGILPDVGVMSVGFREMYLIDAEFKMELVNNRVERPEEYRGALIE 425

Db 357 IPRGIELTRISGHPDQLSMAVSGHVEMILMEAYKYNELINKEEREYESYEGGLVIS 416

Qy 426 PLRGVHENIVLDPSSYVPIMIKYVNGPDTLYRPEKQEC-CGWEAPVGRPRCP 484

Db 417 PLRGVHENIVLDPSSYVPIMIKYVNGPDTLYRPEKQEC-CGWEAPVGRPRCP 470

Qy 485 GFEKTVLERLLELRKVRAMKPKYPPDSPEYRLDDEKOKALYLANASVGMGSGARVY 544

Db 471 GLYKNVLEKIQERKVKKLMER-TIDEYDKRVLDARQRLKMAAAYFYTMGLARVY 529

Qy 545 CRECAKAVTAMGSHLITAINIARLGLKVIYGDTSLFYTYDPKEVNFIKIKELGF 604

Db 530 SKGAEAVTAMGSHLITAINIARLGLKVIYGDTSLFYTYDPKEVNFIKIKELGF 588

Qy 605 EIKLEKVKRLEFTEAKRRVAGLLEGGRIIDVGFPAVRGDMCELAKEVQKVEIYKTS 664

Db 589 EIKDKIKRVKVFTEAKRRVAGLLEGGRIIDVGFPAVRGDMCELAKEVQKVEIYKTS 648

Qy 665 EVKAAVEYKRVYKLEEGKVPLEKLVIMKLSKRLSEYTTAPVVAARMLSAGYRV 724

Db 649 KVEDAIKVVKTVIPDLRRYNEFIEDLIIMKTIDKNLDEYVATAPHVAAKKAAGYLV 708

Qy 725 PGDKIGVIVYKGGRIISQRAPIYFMVNDPSQIDPTVYVDHDIIPALRIIGYFITEKKL 784

Db 709 KGVIGIVYKGGRIISQRAPIYFMVNDPSQIDPTVYVDHDIIPALRIIGYFITEKKL 768

Qy 785 KASATGOKTLPDFLAK 800

Db 769 K---TGVDILSPFK 781

RESULT 4

Q8ZVY2 PRELIMINARY; PRT; 785 AA.

ID Q8ZVY2

AC 08ZVY2

DT 01-MAR-2002 (Tremblrel. 20, Created)

DT 01-MAR-2002 (Tremblrel. 20, last sequence update)

DT 01-JUN-2002 (Tremblrel. 21, last annotation update)

DE DNA-directed DNA polymerase (B3).

GN PAE2109.

OS Pyrobaculum aerophilum.

OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;

OK Thermoproteaceae; Pyrobaculum.

OK NCBI_TaxID=13773;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=IM2 / ATCC 51768 / DSM 7523;

RX PubMed=11792869;

RA Fitzer-Gibson S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,

RA Miller J.H.;

RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum

RT aerophilum."

RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).

DR EMBL; AE009857; AAL63952.1; -.

DR InterPro; IPR002064; DNA_pol_B.

DR InterPro; IPR004578; Pol2.

DR Pfam; PF00136; DNA_pol_B; 1.

DR Pfam; PF03104; DNA_pol_B_exo; 1.

DR PRINTS; PR00106; DNAPOLB.

DR SMART; SM00486; POLBc; 1.

DR TIGRFAMs; TIGR00592; pol2; 1.

DR PROSITE; PS00116; DNA_POLYMERASE_B; 1.

KM DNA-directed DNA polymerase; Complete proteome.

SO SEQUENCE 785 AA; 89547 MW; 8395EAI8C7804F1F CRC64;

Query Match 51.5%; Score 2156; DB 17; Length 785;

Best Local Similarity 53.7%; Pred. No. 1,28-134;

Matches 428; Conservative 111; Mismatches 210; Indels 28; Gaps 9;

[illegible]

RX MEDLINE=20100754; PubMed=10633098;
 RA Kehler M., Antarkitani G.;
 RT "Cloning and Characterization of a Family B DNA Polymerase from the
 RT Hyperthermophilic *Crenarchaeon* Pyrobaculum Islandicum.",
 RL J. Bacteriol. 182:655-663 (2000).
 CC -1- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE = N DIPHOSPHATE
 CC + [DNA] (N).
 CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
 DR EMBL: AF195019; AAF27815.1; -.
 DR HSSP: P56689, ITCO.
 DR InterPro: IPR002064; DNA_pol_B.
 DR InterPro: IPR004578; Pol12.
 DR Pfam: PF03104; DNA_pol_B; 1.
 DR Pfam: PF03104; DNA_pol_B_exo; 1.
 DR PRINTS: PR00106; POLBOLB.
 DR SMART: SM00486; POLBc; 1.
 DR TIGRFAMs: TIGR00592; pol12.1.
 DR TIGRFAMs: TIGR00592; pol12.1.
 DR PROSITE: PS00116; DNA-Binding DNA-directed DNA polymerase.
 DR DNA replication; DNA-Binding DNA-directed DNA polymerase.
 KW PROSITE 785 AA; 89748 MW; 10FB8B66A8F373D CRC64;
 SEQUENCE

Query Match	Similarity	50.8%	Score 2129.5	DB 1	Length 785
Best Local	Similarity	53.4%	Pred. No 76-133		
Matches	Consecutive	111	Mismatches	195	Indels 39; Gaps 13
Qy	9	LDSBYVYWGKPEQVITWGLAENGERYVLIDSFPFRYY--ALLAPGADPQVNAQRAL	65		
Db	8	LDITIAVVGVSPEIRIFGILISSERVLIDSFPFRYYVDCAACEPA-----LKTAL	60		
Qy	66	SRKSPITIGVEDDKRYKFGRRPVLRIRTVLPEAVREYRELYVKNVDGYEDVLEADIRFAM	125		
Db	61	SR-VAEIDVDGIVERRRLGSKKFLKVIAPEDVRLRFAAAMSIPRVSGYEAIDIRFYM	119		
Qy	126	RYLIDHDLPTTWRYVAEAPLENKGFVNDY--YLVKSRPEPLTG--EALAPTKLDLR	181		
Db	120	RYMIDMGVPCSNMVAEEV-----GGRLLGIPYYVVSQ-----WYGDIGCFPS---LK	166		
Qy	182	ILAFDEYVSKSGSPRRPERDPVIAVATKIDGDDVLLPFAEGKDKRQFPIREFEYVSKYDP	241		
Db	167	VMAFDIEYVYNNERSGPDPIRDPVYMLAKITNDGHEVEAAGKDDRGVRAVAPDIIRSYDP	226		
Qy	242	DIIVGNNNNHDMVYLLRRARLIGIKLVYTRRGAETTSYNGHYSVSGRLNDLYHAE	301		
Db	227	DVIVGINSNGDMVYLVERRKAVGPIKVD--RLSNPQQSGVIGMSIVGRANVDLVIIVE	285		
Qy	302	EMPEIKTSGSEAVAEYLGVMKKSERYIINMWEIPDYNDPKKRPILLOYAADVRYATYGL	361		
Db*	286	EMPEIKTGLTVRAVEYGVVKRBERRLPGKHITVEYKDDKNKRLPKTKIVLDDVRSITLGL	345		
Qy	362	AERKILPEAIIQISYVTGLPLDQVAMSVGFRLEWYLIDAAFKMELAVNRVRPEETRYGA	421		
Db	346	ADTLPEFLIQISSVGLPLDQVAAASVGNRVENMILLVYARLGEVAIPRREREVEPYKGA	405		
Qy	422	IVLEPLRGVHENIIVLDFSSMYPNMIIKYNVGPDPTLVRPEKKGEGCGEWAPEYKHFRR	481		
Db	406	IVLEPFGVYEDVIVLDFSSMYNIMMKKNLSPTVYLEPGEPPDPREGVNVVAPYEGHFRFR	465		
Qy	482	CPPGFPTVLRLLLELKRVRAMKMYPPDPSPEYRLIDEGQKALKVLNAAISGYMGSGA	541		
Db	466	SPGFPVQVLKSVLELKAIVREAAKYPDPDSPEKIIDEGQRLAKYMANNAIYGLGVNGA	525		
Qy	542	RMVCGEAKATYANGRLITRLTANTRALGLKLYGPTDSLPTTYDPEKVENFIIKIIEE	601		
Db	526	RMVKEVAVASVTAARAILDQVLEQRRRIGIVVYGTDSLFPFKKHG--VDLTKIYIEEK	584		
Qy	602	LGFEIKLEKVRKRLPFEAKGRVAGLLBDGRIDIVGFEAVRGDMCELAKEVQTVKVEIVL	661		
Db	585	YGDIDIKVDYAKLPTFAKGRVAGLLBDGRIDIVGFEAVRGDMSLAKDQVLKVEIIL	644		
Qy	662	KTSEVWKA-----VVYVKTYIKVELSEGKAPRIKLVIMTYSKRLSEETTEAPHVVAAGML	717		
Db	645	KASDIYEAHRGVICYRIEILERTKYNKKNINDILLIMTTLDEKLEDEKXKPPHYVAAITLK	704		

QY 718 SAGYRSPGDKIGYIVYVGGGRISORAMPYPMVDPQSDIVTYVDHGIIPALRIIGYF 777
 DB 705 RHGRVKGTTIGYIVYVGGGRSERALPYLLDDIKIDIDYIEGIIIPALRIIEVT 764
 QY 778 GITEKJLKASATG--QKTLFDFLA 799
 DB 765 GVKESDLK--TGMERSLIDFLS 785

RESULT 6

Q9P9K4 PRELIMINARY; PRT; 775 AA.
 ID Q9P9K4
 AC Q9P9K4
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE DNA polymerase (EC 2.7.7.7).
 GN DNA POL.
 OS Pyrococcus glycovorans.
 Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 Pyrococcus.
 NCBI_TaxID=74610;
 [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=AL585;
 RA Querrelon J., Cambon M.A., Lesongeur F., Portier P., Barbier G.;
 RT "Thermococcales genes organisation of species belonging to
 thermococcales and phylogenetic implications".
 CC Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 RL -1- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE
 + {DNA} (N).
 CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
 DR EMBL: AJ004334; CAB81809.1; --
 DR HSSP: P56689; ITGO.
 DR InterPro: IPR002064; DNA_pol_B.
 DR InterPro: IPR004578; Pol2.
 DR Pfam: PF00136; DNA_pol_B; 1.
 DR Pfam: PF03104; DNA_pol_B_exo; 1.
 DR PRINTS: PR00106; DNAPOLB.
 DR SMART: SM00486; POLBc; 1.
 DR TIGRFAMs: TIGR00592; pol2; 1.
 DR PROSITE: PS00116; DNA_POLYMERASE_B; UNKNOWN 1.
 KM DNA replication; DNA-binding; DNA-directed DNA polymerase;
 Nucleic acid transferase; Transferase.
 SQ SEQUENCE 775 AA; 90492 MW; 677264920D770C0F CRC64;

Query Match 30.3%; Score 1268; DB 1; Length 775;
 Best Local Similarity 36.5%; Pred. No. 1; 1e-75;
 Matches 304; Conservative 154; Mismatches 278; Indels 98; Gaps 23;

QY 8 VLDSSEYVNGEPVVIIMGLENIGERVVLIDRSPPRYALLAQAADPKQV---AQRIR 63
 DB 2 ILDDVYTEDKQPIIRIRPK-KENGEFKVEYRNFRPYIALLKQDSQIDEVKKTIAERHG 60
 QY 64 ALSRPKSPFIIGVEDDKKRYGPRPRVLRIRITVLPEAVVEYRELVKNGVDVLEADIRF 123
 DB 61 KIYV-----IYVEVKYKKGKFLGRPIEVKGLYFEHPQDVAIIDKIREHAAVVDIEYDIPF 116
 QY 124 AMRYLIDHDLFPFTWYRVEAPELENKMGFRVQKYLKVSREPPLGEMALAPTKLPDLRL 183
 DB 117 AKRYLIDKGLI-----PMEGD-----ELKTL 138
 QY 184 AFDIEVSKQSPRPERDPVIVIAKTDSDVLT-----FLAEGDDKRPREPVEY 235
 DB 139 AFDIETLVHEEB-EPAKPIIMISADEGAKVITWKVDDPYVEVSSEREMIKRPLKV 197
 QY 236 VKRYDPIIIVYNNHFMPIILRRARILGILDVTRVGAEPPTSVAHG---VSPGR 292
 DB 198 IREKDPDIIIVYNGDSFPLPIVYGAELGILKPLGRD-GSEPKQRIQDMDTAIVEIKRI 256
 QY 293 NVLDYVAEMPEIKIKSLAEVAVYLGVWKSSEVVIINMWEIPDYMDPKRPLLLQYAR 352
 DB 257 HPDLVHVIRRTINLPYTLAEVYEA-EGKPEKRYAH--EIAEMETGKGLERVAKYSM 313

QY 353 DDVATYGLAKTILPFAIOISYVTGLPLDQVGMVSGFRLWLTIRAPFMKELVNRVE 412
 DB 314 EDKAVTYELGREFFPMDAQSLRVGQPLMDVSRSSSGNVLTKKAAERNELANEKPD 373
 QY 413 RPE-----ETVGAIVLEPLRGVHENIAVLDFSSMYPMIMKYVNGPDTLVPRGECSE 466
 DB 374 EREYERLRRESVAGYVKEPEKGLMEGLVSLDFRSLYPSIIITHNVSPDLTNE----- 427
 QY 467 CGCWE---APEVGRFRRCPPGFKTVLERLLELRKRVRAEMKKYPPDSEYRLDEROK 523
 DB 428 -GCMEXYVAPEVGAKECKPFGFIPSLKRLDLDEROIKRMRKA-SKDPLEKMLDYRGR 485
 QY 524 ALTVLNASVGMVNGSGARVCECAKATVAMRHLIR-TAINIAKTLGKVIYGGTDS 562
 DB 486 AIKILANSYGYGVAKARVYCKEASVYAKREYIEFVRKLEKFGKVLVITDGL 545
 QY 583 FVTY---DPEKVB---NFIKIKEIG--FEIKLEKVVYRLPFEAKKRYAGLLEDGRI 633
 DB 546 YATIPAKPEIRKALKEFEYINAKLPGLLEIEYGFYRGFFV-TKKYALIDEGKI 604
 QY 634 DIVGFEAVRGDMCELAKEVQTKVEIVLKTSEVYKAVEYVKIVKLEBEGVPIEKLVIW 693
 DB 605 ITRGLIETVRDMSEIAKETQAKVLEALIKRGVBEAVKIVKEVTEKLSKYEIPPEKLV 664
 QY 694 KTLASKLEETTAAPRVYNAKRLSLAGYRSPDKIGYIVYVGGGRISORAMPYPMV-- 751
 DB 665 EQITRPLHEKALGPHVAVKRLAARGVVRKPMVIGYIYLRDGGISKKA---ILAEER 721
 QY 752 DP--SQIDVYVVDHGIIPALRIIGYFITEKTLKASATGQKTLPLDKSKK 803
 DB 722 DPKHKYDAIYIENQVLPVAILLEAFGRKEDLMQTKQIGLAWLVKCK 775

RESULT 7

Q9HH06 PRELIMINARY; PRT; 775 AA.
 ID Q9HH06
 AC Q9HH06
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE DNA polymerase.
 GN POL.
 OS Pyrococcus glycovorans.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OX Pyrococcus.
 NCBI_TaxID=74610;
 RC SEQUENCE FROM N.A.
 RP STRAIN=AL646;
 RA Querrelon J.J.E., Cambon M.A., Lesongeur F., Barbier G.;
 RT "Thermococcales taxonomy and phylogeny based on the comparative use of
 16S rDNA, 16S-23S rDNA intergenic spacer and family B DNA polymerase
 genes".
 CC Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 RL -1- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE = N DIPHOSPHATE
 + {DNA} (N).
 CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
 DR EMBL: AJ250335; CAC12849.1; --
 DR HSSP: P56689; ITGO.
 DR InterPro: IPR002064; DNA_pol_B.
 DR InterPro: IPR004578; Pol2.
 DR Pfam: PF00136; DNA_pol_B; 1.
 DR Pfam: PF03104; DNA_pol_B_exo; 1.
 DR PRINTS: PR00106; DNAPOLB.
 DR SMART: SM00486; POLBc; 1.
 DR TIGRFAMs: TIGR00592; pol2; 1.
 DR PROSITE: PS00116; DNA_POLYMERASE_B; UNKNOWN 1.
 KM DNA replication; DNA-binding; DNA-directed DNA polymerase.
 SQ SEQUENCE 775 AA; 90417 MW; 633F8392EDBEEC73 CRC64;

Query Match 30.2%; Score 1264; DB 1; Length 775;
 Best Local Similarity 36.5%; Pred. No. 2e-75;

Matches 304; Conservative 154; Mismatches 278; Indels 98; Gaps 23;

QY 8 VLDSSEYVGEKPEOVIIWGIAGNGERVVLIDRSPRPYFVALLAPGDPKQV----AQRIR 63
 DB 2 ILADVITDEGKPIIRIFK-KENGEKVEYDRNRRPYIYALLDDQIDQIDVKKIRNKHG 60
 QY 64 ALRPRKPIIGEDDCKRYGPRRRLRRTVTEPAVREKLVKNDGVEYLAADIRF 123
 DB 61 KIVR-----IVDEKVKKKKFKGPRRIEWMKLYTEHNDPAIRDKIREPAVVIDEYDIPF 116
 QY 124 AMRYLIDHDLPFTWTRVNAEPLKNGKRRVDKYLKVSRRPELYGALAPTKLPDLRI 183
 DB 117 AKRYLIDKGLI-----PNEGQ-----BELKL 138
 QY 184 AFIDIEVYSKQSGPRPERDPIYIAVKTDGDEVL-----FLAEGKDRKPIREPV 235
 DB 139 AFIDETLYHGE-EFAGPIIMISYADEGAKVITWKKVDLPYVEVSSEREMIRPKV 197
 QY 236 VKRYDPDIIVGNNHEDWPFYLLRRARILGILKDVTRVGAEPFTSVHG--VSVPRL 292
 DB 198 IREKDPVVIITVNGDSFDLPYLKQAEKLGILPLGHD-GSEPMQRLGDMTAVEIKGRI 256
 QY 293 NVLDYDAEEMPEIKIKISLEEAVALGWKKSERYIINMWEIPDYMDPKQPLLOYAR 352
 DB 257 HFDLYHYIRRTINLPYTLAEVYEA-FCRKEKVVAN-EIAMEWETGKLEBAVAKSM 313
 QY 353 DDVATYGLAEKILPFAIOLSVYTGPLPDYOGASVGFRLBWLIRAAFKMELVPRVE 412
 DB 314 EDKATYELGREFPPEMAQSLRVLGQPLMDVSRSTGNLVEMWLAKVERBELAPNKP 373
 QY 413 RPE-----ETVGAIVLEPLRGVHENIAVLDPSSMYPNIMIKYNGPTLVPRBEKGE 466
 DB 374 EREVERLRLESYAGVYKBEKMGVLSDFRSLPISIIITHNSPDLNRE-----427
 QY 467 CGCWE---ABEVKRRRCRCPGPFKTVLERLLEKRVAKEMKKYPPDSEPRLLDEKQ 523
 DB 428 -GCEYDAVEBKFKCKDPGPIPSLRRLRLLDEGQEIKKRMA-SKDPLEKMLDYQOR 485
 QY 524 ALKVLNANSYGMWGSAGARVYCECAKAVTAWGNHLIR-TAINIARKGLKVIYGDTS 582
 DB 486 AIKTLANSYGYGYGAARVYKCEBCESEVYTAGREYLEFVRKLEBKFPFKLYIDTGL 545
 QY 583 FVTV--DPEKVE---NFKIKIKELG--FEIKLEKVRQLPFTBAKRRYAGLLDEGRI 633
 DB 546 YATIPGAKPEEIKRKALEFVEYINAKLPGLLEBEYEGFYRGEFV-TKKVYALIDEBGKI 604
 QY 634 DIVGFEAVRGDMCELAKEVOTKVEIVLKTSSVNAVAEYVRKIVKELEGKPIREKVI 693
 DB 605 ITRGLEIVRRDMSEIAKETQAKVLEILKHGVNEAVKIVKYTEKLSKYEIPREKVI 664
 QY 694 KTLSEKRLSEYTTTEAPHVNAKMLLSAGRVSPGKIGVIVKGGSRISORAMPYMW--751
 DB 665 EQITRPLHEKKAIGPHVAAKGLAARGVKVRPGWVIGIYLRDDGISRGA---ILAEF 721
 QY 752 DP--SOLDVTVYVDDHOLIPALRIIGYIGITEKLSAKSAQOTLIDPLAKSK 803
 DB 722 DPKHAKDAEYIENOVLPRAVIRLLEAFGRKGDRLKOKTKOGLTAMLVKK 775

RESULT 8
 Q8TWJ5 PRELIMINARY; PRT; 830 AA.
 AC Q8TWJ5; 01-JUN-2002 (Tremblrel. 21, Created)
 DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE B family DNA polymerase.
 GN POLB OR MK1039.
 OS Methanopyrus kandleri.
 OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
 OC Methanopyrus.
 NCBI_TaxID=2320;
 RN [1]

RP SEQUENCE FROM N.A.
 RC SRRAIN-AV19 / DSM 6324 / JCM 9639;
 RX MEDLINE-21927647; PubMed-11930014;
 RA Steerey A.I., Mezheveya K.V., Makarova K.S., Polushin N.N.,
 RA Shcherbina O.V., Shakhova V.V., Belova G.I., Atavind L.,
 RA Natsale D.A., Kogozin I.B., Jatusov R.U., Wolf Y.I., Stelcer K.O.,
 RA Malyshe A.G., Koonin E.V., Kozaykin S.A.;
 RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19
 RT and monophyly of archaeal methanogens";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649 (2002).
 DR EMBL; AE010392; AA00252.1; --
 KW Complete Proteome.
 SQ SRODNCE 830 AA; 94561 MW; 094B155D724EE47 CRC64;

Query Match 29.1%; Score 1220; DB 17; Length 830;
 Best local similarity 35.0%; Pred. No. 1.9e-72;
 Matches 299; Conservative 156; Mismatches 285; Indels 114; Gaps 27;

QY 19 EPQVIIWGIAGNGERVVLIDRSPRPYFVALLAPGDPKQVQRIRALSRKSPPIIGEDD 78
 DB 14 EPDVLIVGRREDQNPALVLVKGRFPYFVLEVEDGDPSEV-ERL-----SGVVEVEV 65
 QY 79 --KRYKGRPRRVLRIITVLPVAVREYRELKNDGVEYLAADIRPAMEYLLDHDLPF 136
 DB 66 LLEHPIGDRBELRIYVATPKVYPLRKQVKKLQVGEYEAADIPVRRAVDNLNPPA 125
 QY 137 TWRYDAEPLN-----KMGPRVDKVIYVSRPRELYGALAPTKLPRLIADIV 189
 DB 126 S--EVDVSDLDGMSGLPAPFADVEDARELHRRY-----TEDLVASDLSEV 173
 QY 190 YSKQSGP-RPERDPIYI--AKTIDGDEVFLA--BGKD-----DRK 227
 DB 174 LMEPTTIGASGPIIISPRYSTPDEGRNRNVITWKEDESFVGVETEVIVCSESA 233
 QY 228 PIREVEYKRYDPIIVGNNHEDWPFYLLRRARILGILKDVTRVGAEPFTSVHG--284
 DB 234 ALRPFDEBRRDDVVFYTNGBDFLPYLQHRACKLIDIVSLPARAGRGIIILGGG 293
 QY 285 --HVSVPGLNVDLYDAEEMPEIKIKISLE-EVAEYLGWKKSERVLIINMWEIPDYMDP 341
 DB 294 RYASDIFGAAHYDLVHARAKLKEPFTLEEAUVKVLVEKEBELA---DINEAM---346
 QY 342 KKRPL--LLQYARDDVATYGLAEKILPFAIOLSVYTGPLPDYOGASVGFRLBWLIR 399
 DB 347 KRGNDELKRYASADAYTLELGLAEQVELSTYTRPLPDATRRSPQLAEW--RA 403
 QY 400 AFGKKE---LVNPNVREPE-----ETVGAIVLEPLRGVHENIAVLDPSSMYPNIMIKY 450
 DB 404 IYANQEDILVNNKPTDEYKRRRRAKYGALPPEPELGHNVCVDPAFLPNVMAH 463
 QY 451 NVGPDIVLRPBEKCGCWE-----ABEVKRRRCRCPGPFKTVLERL 495
 DB 464 NISDFP-----DCDCPRVTVVEVDPLDPAVADVGHKCKRKRKFPRRLVEGL 515
 QY 496 ELKRVRAEMKKYPPDS--PEYRLDEROKALVLANASYGMWGSAGARVYCECAKAVT 553
 DB 516 ERRELKRLRLRLDTESHPEAKILDVROAAKVLANSYGYWGMANARFCEBCESEST 575
 QY 554 AMGRHLIRTAIINIA-RKLGKVIYGDTSLPFTYDP-----EKVENFKIKIKELGPE 605
 DB 576 AMGRYIISEVRRIAEKYGGLKVVYGDTSLPFKLPDADLEETIRKVEFLKEVNGSLPPE 635
 QY 606 IKLEKVRQLPFTBAKRRYAGLLDEGRIDIVGFEAVRGDMCELAKEVOTKVEIVLKTSE 665
 DB 636 LELEDYKRIILFV-TKKVYAGYEDKIVTKGLIFVRBDMAPIRAFQORVLKRIADND 694
 QY 666 VNAVAEYVRKIVKELEGKPIREKVIYMTLSKRLSEYTTTEAPHVNAKML-SAGRVRS 724
 DB 695 PEALKEHEVLEBRLKSGDVVIDELAVTSQLTKKSESEVYQGPVRAALRLAHGVEPE 754
 QY 725 PGDKIGYIVVGGSRISORAMPYMKDP--SOLDVTVYVDDHOLIPALRIIGYIGITEK 783
 DB 755 PGIYRVYIVVPGSVSDKAVPELVREBGEKPDVITLIEHOLIPAVERIMRAIGISRQ 814

QY 784 LKASATGOKTLFDF 797
DB 815 IGVETASOKTIDOF 828

RESULT 9

Q9HH98 PRELIMINARY; PRT; 759 AA.

DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, last annotation update)
DE DNA polymerase (Fragment).

OS Pyrococcus sp. (strain ST700).
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX [1] NCBI_Taxid=69015;

SEQUENCE FROM N.A.

STRAIN=ST700;
RA Querellou J.-J.E., Cambon M.A., Lesongeur F., Barbier G.;
RT "Thermococcales taxonomy and phylogeny based on the comparative use of
RT 16S rDNA, 16S-23S rDNA intergenic spacer and family B DNA polymerase
RT gene.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE = N DIPHOSPHATE
CC + (DNA) (N).

CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.

DR EMBL; AJ250332; CAC12847.1; -.

DR HSSP; P56689; IRGO.

DR InterPro: IPR002064; DNA_pol_B.

DR InterPro: IPR004578; GPCR_Rhodopsn.

DR Pfam; PF00136; DNA_pol_B; 1.

DR PRINTS; PRO0106; DNAPOLB.

DR SMART; SM00486; POLBc.1.

DR TIGRfams; TIGR00592; Pol2.1.

DR PROSITE; PS00116; DNA_POLYMERASE_B; UNKNOWN_1.

DR POSITIVE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.

KM DNA replication; DNA-binding; DNA-directed DNA polymerase.

FT NON TER 759

SO SEQUENCE 759 AA; 88431 MM; A8459B6AAD8F3A3 CRC64;

Query Match 28.8%; Score 1208; DB 1; Length 759;

Best Local Similarity 35.4%; Pred. No. 1e-71;
Matches 289; Conservative 156; Mismatches 273; Indels 98; Gaps 21;

QY 8 VLDSYEVNKGEPVVIIMGIANENGRVLLIRSPRPYYALLARCA--DRQV-AQIR 63
DB 2 ILDDVYTNCKPIIRLTK-KENGFKVEYRNRPYIYALLKDSADIVRKITSEHG 60
QY 64 ALSRPKSIILIGVEDDKRYFGPRPRVLRITVLPFAVREYRELVKNDVGEVLEADIRF 123
DB 61 KVA-----VIVEKSKSKFGRPIEVWGLYEHPODVAIIDKREHNAVVIDEYIDPF 116
QY 124 AMRYLIDHDLFPFTWYRVEAPLENKMGFRVQVYLVKSREPLYGALAPTKLPDLRIL 183
DB 117 AKRYLIDKGLI-----PMEGN-----EELGFL 138
QY 184 AFDIEVSKOSPRRPRDPVIVIAVKTDGDEVL-----FLAGKDDKPIREPREY 235
DB 139 AVDIETLYHEE-BFGKPIIMISADBEAGKVITWKIDPYVEEVANEREMIKRLVRI 197
QY 236 VKRSDPDIIIVGNNHFWPILIRARILGKLDVTRVGAAPTSSVGH---VSVPRL 292
DB 198 IREKDPDIIIVNGDNFPFLLKRAEKGILKPLGR-NSEPKQRLGSLVAIEIKRI 256
QY 293 NVLDYVAEMPEIKSLAEVAEYLVGWKSERVIINMWEIPDVMDPKRPLLYAR 352
DB 257 HFDFPPIVIRRTINLPYTLRTVEAI-FGKPKKVPYH-EIAEMWEGKGLERVAKYSM 313

QY 353 DVRAATYGLAKTILPRAIOHSYVTGLPLDOVGAMSVGRLEWYLIRAPFKMLVPRVE 412
DB 314 EDKAVTYELKEFFPMEAOIARLVGGPMDVSSSTGNLWELRLRAVYERNEIAPKPD 373
QY 413 RPE-----ETRGALVLEPDRGVHENIAVLDFSSMYPMIKVNPDPDLVRPEKCGE 466
DB 374 EKEVEKRLRESYEGGYKEPEKGLMEGIVSLDFRSLPSIITHNNSPDLNREG----- 428
QY 467 CGCW-EAPEVKHRRRCPPGPFVTLERLLERKRYAENKKYPPDPSPEYLLDERQAL 525
DB 429 CGKDEAPEVGHFCKDPFGFIPSLGLDILLEEROKIKRKE-SKDIPEKLLDYQRAI 487
QY 526 KVLANASYGMSGAMWYRECAKATANGRLIRAIINIAKLGKVIYGTDSIFVT 585
DB 488 KILANSFYGYGAKAWYCEKCAESVTANGROYIELVRLEBERGKVLITDGLYAT 547
QY 586 YDPEKVENFKIIEELGF-----EIKLEKVVRLPFTFAKKRYAGLLEDGR 633
DB 548 IPGEK--NMEIKRKALFVNYINSKLPGLLELEYBGFYTRGFV-TKKYALIDEGKI 604
QY 634 DIVGFEAVRDWCELAKEVQTVKVEIVLKTSEVNAVEYRKIVKELEBGKVPTEKLV 693
DB 605 ITGLEIVRDMSEIAKETQAVLEAILKGNVEAVKIVKYTEKLSNVEIPEKVIY 664
QY 694 KTLSKLREYTTAPHYVAKRLSAGYRVSBDKIGYIVKGGRIQAMFYVVKD- 752
DB 665 EQTRIPNEKAIQPHVAVKRLAKGIRKIPGVIGYVLLRDPGISKBA---IAIEEF 721
QY 753 ---PSQIDVTVYVDHQTIRLIPALRILIGFYITEKKL 785
DB 722 DGKHKTDAYTIENQVLPAVERILIRAFYKEDLR 757

RESULT 10

Q8TSB3 PRELIMINARY; PRT; 937 AA.

AC Q8TSB3;

DT 01-JUN-2002 (TrEMBLrel. 21, Created)

DT 01-JUN-2002 (TrEMBLrel. 21, last sequence update)

DE DNA-directed DNA polymerase.

GN MA0885.

OS Methanosarcina acetivorans.

OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;

OC Methanosarcinaceae; Methanosarcina.

OX NCBI_Taxid=2214;

RP SEQUENCE FROM N.A.

RC STRAIN=C2A / ATCC 35395 / DSM 2834;

RX MEDLINE=21929760; PubMed=11932238;

RA Galagan J.E., Nussbaum C., Roy A., Endrizzi M.G., Macdonald P.,

RA Fitzhugh W., Calvo S., Engels R., Smirnov S., Alnoor D., Brown A.,

RA Allen N., Naylor J., Stange-Thomann N., DeArjano K., Johnson R.,

RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,

RA Zimmer A., Barber R.D., Camm I., Graham D.E., Grahame D.A., Guse A.M.,

RA Hedderich R., Ingram-Smith C., Kuettnet H.C., Kzycki J.A.,

RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,

RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,

RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,

RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,

RA Metcalf W.W., Birren B.;

RT "The genome of Methanosarcina acetivorans reveals extensive metabolic

RT and physiological diversity.";

RL Genome Res. 12:532-542 (2002).

DR EMBL; AB010753; AAM04324.1; -.

KM DNA-directed DNA polymerase; Complete proteome.

SO SEQUENCE 937 AA; 105600 MM; 4ADAB7F53AD59B6 CRC64;

Query Match 20.4%; Score 852.5; DB 17; Length 937;

Best Local Similarity 31.4%; Pred. No. 5.7e-48;
Matches 294; Conservative 132; Mismatches 324; Indels 185; Gaps 43;

```

QY 6 FTVDDSYEVVKE-POVITWIGIAENGERVVLIDRSPPRYFALLAAGADPKQVQNRIR- 63
DB 5 FQILDADYEVVNSGCPVIRLFGAGADKSCVCPVDFPFPYTY--LKASGDIAVAVRLKD 62
QY 64 ALSRKSPRIIGVEDDKKRYGGRPRVLRRLTLPJANRER-ELVK-----NVDGYEDVL 117
DB 63 TPBOYKKEVEKEFEPVIGTCKEMKVTTLTLPDQVPELIDELIKRDVLAAGQMOY 122
QY 118 EADIRFAMRYLIDHDLFFPFWYRVEAPLE--NKGK-----FRVDKYVLK--S 162
DB 123 ESDILFNRRLIDALGGMVASBGPVDPVRYAGASANRSCENFADSAVLASGLK 182
QY 163 RPEPLYGELALPTKLPDLRIALADIEVYSKGS--PRPERDVYIAVKTD----- 211
DB 183 RVENL--ALAP-----LKYALADIECLPLDGMSPDPVSIIMISFSEPEYKHKTLI 234
QY 212 -----DGEVLFLAEGKDRKPIREFVEYVYKYPDIIYGVNNNHFDMPYLLRBAR 262
DB 235 LLAKEPAGMDGD-----VLSCMDETEMANKFEIICEYDPDIYAGYNNHODEDIPYTERVK 290
QY 263 ILGILKLDVTRVGAEPITTSYHGVSVP-----GRLANVDLYDAEEMPEIK 307
DB 291 ALVAK-----GETINSVVGSDSPICGRKFGILTRTEMKGRVYVDALPLVRAAFGLK 342
QY 308 IKSLEEVAVLYGVKSKSERIIMMEIDPYMDP--KXRPILLOYADDPVATYGLA--EK 364
DB 343 QYTLRAVSKEL--LSREKLDVPRLEMEHNSGDKRKRVYARDSLSALVELYLRL 399
QY 365 ILFPALQUSYTTGLPLDV--GANSVGRLEWYILRAAFKMEYVNRVERPE----- 415
DB 400 LDKYTLAQSGLLOEIVDGOQTSW--VELTLRL-EGEKD--RVLIPKQDELS 451
QY 416 -ETY-----RGALVEPLRGVHENIAVLDFSGSNYIMIKYVNGDPTLV--RPEKCC- 464
DB 452 AERYDMSDLKGGVLEPKKGLLENVLLIDYKSLPTTMAHMLCTYTVVTRDRPDKTI. 511
QY 465 -----GEGCWAEAEVGRFRRCPPGFFVTYLERLLERLKRVAEMKKYPPDPSPERYLID 519
DB 512 KPPSCGE--FVPEPV--FR--GIYPSIIEIDLNNKRGDTKMKMR--TSDENRHYLD 560
QY 520 ERQKALKVLANASYGYMGSGARWYCRECAKAVTAMGR--HLIRTAIINARKG----- 571
DB 561 ATQALIKILNSFYGYSGYARALYSLTLANAVTSFGRSNLTMBDLINGRIGKIVLANS 620
QY 572 -----LKVYGTDSLFPV-----TYDPEKV--ENFIKIIXEEL 602
DB 621 AALLBEAGKLSPODRIVELSVAYGDTDSVFVHCACKODLSLEVSIVGNRLSEIVASL 680
QY 603 --GPEIKLEKYKELFTTEAKKRYA-GILE-----DGRIDYGFVAVGDMCELAKEXQ 653
DB 681 PDPMELEFESVAKRALLI-AKKRYALMLFEPNNSGEMKLVKGMETVARDWCBLISITL 739
QY 654 TKVVEIVLKTSEVNAKAVEYRKI--VSELEBGKP--TEKLVIMKTLKRLSEYTTENP 708
DB 740 NNVLEFVLISGDVDAVEHAKVSDVRNLDPGAGITTELVITRLTRKADSKYNQOP 799
QY 709 HVAAKRMLS-AGRYVSPGDKIGYIVIKVGGRIISORAW--PYFWVKPSPQIDYTYVYDQI 766
DB 800 HLTVAENLKRRGTGMPISIGRILPFPVITACKGLFVDRADDPVYRENNVPIDYDYYVKQI 859
QY 767 IPAAALRIVGEGITEKKLKASATGOKTLPDPLAKK 801
DB 860 LPVVERILEVGVVMSLSLDFDAK-QKGLDFEVKK 893

```

RESULT 11

ID 096H98 PRELIMINARY; PRT: 1107 AA.

AC 096H98: 01-DEC-2001 (T-EMBLrel. 19, Created)

DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)

DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)

DE Polymease (DNA directed), delta 1, catalytic subunit

```

DE (125KD)
OS Homo sapiens (human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN (1)
RC SEQUENCE FROM N.A.
PC TISSUE=Lymph;
RA Strausberg R.;
RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: N DIOXYNUCLEOSIDE TRIPHOSPHATE = N DIPHOSPHATE
CC + (DNA) (N).
CC -1- MISCELLANEOUS: IN EUKARYOTES THERE ARE FIVE DNA POLYMERASES:
CC ALPHA, BETA, GAMMA, DELTA, AND Epsilon WHICH ARE RESPONSIBLE FOR
CC DIFFERENT REACTIONS OF DNA SYNTHESIS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
DR EMBL: BC008800; AA008600.1; -.
DR InterPro: IPR002064; DNA_pol_B.
DR pfam: PF00136; DNA_pol_B: 1.
DR TIGRfam: TIGR00592; pol2.1.
DR PROSITE: PS00116; DNA_Polymerase_B: UNKNOWN 1.
DR DNA replication, DNA-binding, DNA-directed DNA polymerase.
KY SEQUENCE 1107 AA; 123611 MW; 108A832719165DBA CRC64;
SQ
Query Match 18.2%; Score 760.5; DB 4; Length 1107;
Best Local Similarity 28.7%; Pred. No. 9.4e-42;
Matches 255; Conservative 144; Mismatches 320; Indels 169; Gaps 33;
QY 17 GKEPQVITWIGIAENGERVVLIDRSPPRYFALLAAGADPKQVQNRIR--ALSREK--- 69
DB 120 GSVPLVIRAGVYDSEFSVSCYCHINGFAPYFPYAPGFPBEMHGLQREMLAISRSRGC 179
QY 70 -----SPIIGVEDDKR-----YFGR--PRVRLIRVTLPEAAREVRELVKAVDDEVY- 116
DB 180 RELTGAVALVALVCSRESMEGYGHGSPFLRITVALRLVA PARRLLEGIRVAGLCTP 239
QY 117 -----LEADIRFAMRYLIDHDLFFPFWYRVEAPLEENKGFVVDKVVY-----YKSRP 164
DB 240 SPAPYAVANDVEIRNMDVTDIVGCMNLELPACKVALRLKEKATCGQLEADVLSVDVSHP 299
QY 165 EPLYGALAPTLDPDLRIALADIEVYSKGS-SPRERPVVI--AVKDDSDVLTFLTA 220
DB 300 PEPMPKRIAP-----LWLSFDIFCAKRGKIPPEBRDPVQISLGLRGEPEPFLRLA 354
QY 221 -----EKODRKPIREFVEYVYKYPDIIYGVNNNHFDMPYLLRRA-- 261
DB 355 LTRPCAPILKAKVOSYKESDL--LQAWSTIRIMDDPVITGVNIONFPLPYLISRAOT 412
QY 262 -----RLIGIKLDV-----TRVGAEPITTSYHGVSVGRNLNVDLYDAEEMPE 305
DB 413 LKQVTFPLGRVAGLCSINRDSFOSKQGRDRTV--SVNMGROYMDLQVL--LNE 466
QY 306 IKIKS--LEEVA--EYLGVKMSKSERVLIINMWEIDPYMDPKRPLLAQYARDVR--ATY 359
DB 467 YKLRSTYLANAVSFHFLGQKEDVOHSI--LTD-----LQNGDQTRRLAVY 511
QY 360 GLAEKILTP-----AIOQSYVTGLPLDQVANSVGRFLEWYILRAAFKMEYLP 408
DB 512 CLADAVLPLRLERLMLVANAEMARVGPVLSYLSGQGVKVVSGLQANHEGLMWP 571
QY 409 NRVREEREYRGATVLEELRGVHE-NIAVLDSGSNYIMIKYVNGDPTLVRG--EKGS 465
DB 572 VVKSQEGEDYTGATVIEPLKGYDVIALDSSLYPSIIMAHMLCTTILIRGTQKQKG 631
QY 466 EGCWAEAEVGRFRRCPPG--FFKT-----VLERLLERLKRVAEMKKYPPDPSE 514
DB 632 -----LTBOQFLRTGTGBEFVKTSVKGLLPOLLENLISARKRKAELAK-ETDPLR 682
QY 515 YRLDERQKALKVLANASYGYMGSGARWYCRECAKAVTANGRLI-----RTAIN 565
DB 683 RQVLDERQALKVANSVYGTGAQVQKLPCLIEISQSYTGVGROMIEKQKQVESKYTVE 742

```

```

QY 566 IARKGLKVIYGDPTDLSFVTYDPEKVENFIKIKEELGF-----EIKLEKYYKRLF 616
DB 743 NGYISTSAKVYIGDDTSVWCRGVSVAEAMALGREAADVSGHFPSPRLREKYYFP-Y 801
QY 617 FTEAKRYVAGLL-----EDGRIDIVGEFAVRGWCMLAKEVQTVKEIVLKTSEVNKAV 670
DB 802 LLISKRYVAGLLFSSRPDHRMCKGLEAVRRDNCPLVANLVTSARLLRIDNDPEGAV 861
QY 671 EYAKRYVELEEGKVPLEKVIYWKLSKRLSEYTTAPRHVAAKRM--LSAGYRVSFGDK 728
DB 862 AAHADVSDLCNRRIDISQLVITTELTPAADYAGKQAHVELAEMRRKRDGSAFSLDR 921
QY 729 IGYIVY---KGGGRISQRAWPFYMKDPSQIDVYVYVHOIIPALRI 773
DB 922 VPYIISAKGVAAVYMKSEDPFLVLEHSLPIDTYTLEQLAKPLRLI 969

RESULT 12
PRT: 1105 AA.
091VTO PRELIMINARY;
01-DEC-2001 (Tremblrel, 19, Created)
01-DEC-2001 (Tremblrel, 19, Last sequence update)
01-JUN-2002 (Tremblrel, 21, Last annotation update)
DE DNA polymerase delta 1, catalytic domain.
GN POLDI.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NX NCBI_Taxid=10090;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=BREAST TUMOR;
RA Strausberg R.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATES = N DIPHOSPHATE
+ {DNA} (N).
CC -1- MISCELLANEOUS: IN EUKARYOTES THERE ARE FIVE DNA POLYMERASES:
ALPHA, BETA, GAMMA, DELTA, AND EPSILON WHICH ARE RESPONSIBLE FOR
DIFFERENT REACTIONS OF DNA SYNTHESIS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
DR EMBL: BC009128; AA009128.1; -.
DR MGD: MG1:97741; PolDI
DR InterPro: IPR002064; DNA_pol_B.
DR InterPro: IPR004578; Pol2.
DR Pfam: PF03104; DNA_pol_B_exo; 1.
DR TIGRFAMs: TIGR00592; Pol2; 1.
PROSITE: PS00116; DNA POLYMERASE B; UNKNOWN 1.
KW DNA replication; DNA-binding; DNA-directed DNA polymerase.
SQ SEQUENCE 1105 AA; 123744 MW; 20323690DD472406 CRC64;

Query Match 17.8%; Score 743.5; DB 11; Length 1105;
Best Local Similarity 27.9%; Pred. No. 1.3e-40;
Matches 246; Conservative 153; Mismatches 319; Indels 165; Gaps 33;

QY 20 PQVITWIGIAENGERVYLIDRSFRPYVALLPAGADPKVQVQIR---ALSRRK----- 69
DB 121 PILRAFGVTDEGFSVCCHIQGFAPFYTPAPPGFGEHLSLQGLTMAAISRDGCKGKEL 180
QY 70 --SPVIGEDDKRK---YFGR-PRRVLRIRTVLPEAVREVELVK---NVDGVD----- 115
DB 181 SGPAVLALELCSRESMFCYHGGSPPLRITLALPRMAAPARLLLEQGVRRVFGTGPFA 240
QY 116 VLEADIRFAMRYLLIDHDLPFTW-----YREAPELENKGFVVDKYYL-VKSRPPL 167
DB 241 PYEAVNDPEIRFMVDADIVGCNMWLEPAGKYVRAREKATLCQLEVDVIMSGVISHPEEG 300
QY 168 YGEALAPFKLPDLRIALVDIEVYSKQ-SPPRRDPVITV---AVKTDGGEVLFIA--- 220
DB 301 OMORIAP-----LRVTSFDIECAGRKGIFFEPERDPVIGICSLGRMGEPEPFLALTL 355

```

```

QY 221 -----EGKDDRKPIREFVSVRYDPODIIVGVNNNHDPMYLARRA----- 261
DB 356 RPAAPILGAVQSYEREDL--LQAMADFLAMPDPVITGYNIONFPLPYLISRAQALKV 413
QY 262 -----RIIGIKLDV-----TRVGAEPPTSNGHGVSPGRILNVDLYDAEEMPEIKI 308
DB 414 DRFPPLRGVTLGNSINRDSFSGRQVGRDSKY-----ISWVGKQVMQMLQVLRHNTLS 469
QY 309 KSLEVA-EYLGVMKSERVIINMWEIPDYWDPKRRPLLQVARDVR--ATYGLAEK 364
DB 470 YTLNAVGFHFLGEOQEDVQHSI-----ITD-----LONGNEQTRRLRAYCYCKDA 514
QY 365 ILPEA-----IQLSVYTGCLPLDYGANSVGRLEMYLIRAFKKEIVPRVER 413
DB 515 FLPLRLRLRLMLVNNVEMARVTVGPVPLGYLLTRGOQVAVSOLLRQMRGGLMPVVKTE 574
QY 414 PEETRGAVILEPLRGVHE-NIATLDFSSMYNPMIKYNGVPTLVLPFG--EKCGEGCW 470
DB 575 GGEDYTGATVIEPLKGYVDVPATLDFSSISYPSIMAHNLCYTTILRPGAQKLG----- 629
QY 471 EAPF-----YKRRFRCPGPFKTVLERLLELRKRVRAEMKRPDSPEYRLD 519
DB 630 LKPDFFIKTPTGDEFVSVYRK--GLLPOLLEVLASARRAKAELAQ-ETDPLARQVLD 685
QY 520 ERQALVILNANASYGVNGMGARVYCECAKATYAMGRHLIRPAINARK-----L 570
DB 686 GRQALTKVANSVYGFYGAOVGKPLCLEISQSVTFGRQMIETKQOLVESKRYTVENGYDA 745
QY 571 GLKVIYGDPTDLSFVTYDPEKVENFIKIKEELGF-----EIKLEKYYKRLFTEAK 621
DB 746 NAKVYIGDDTSVWCRGVSVAEAMSLGRBAAMVSSHFPSPRLREKYYFP-YLLISK 804
QY 622 KRYVAGLLDGR-----IDIVGEFAVRGWCMLAKEVQTVKEIVLKTSEVNNAVEYVRK 675
DB 805 KRYVAGLLFSSRSDAHDMDCKGLEAVRRDNCPLVANLVTSARLLRIDNDPEGAVAAK 864
QY 676 IYVELESGKVPLEKVIYWKLSKRLSEYTTAPRHVAAKRM--LSAGYRVSFGDKIGVI 733
DB 865 VTSDLNCRIDISQLVITTELTPAADYAGKQAHVELAEMRRKRDGSAFSLDRGVYVI 924
QY 734 V---KGGGRISQRAWPFYMKDPSQIDVYVYVHOIIPALRI 773
DB 925 IGAAGVAAVYMKSEDPFLVLEHSLPIDTYTLEQLAKPLRLI 967

RESULT 13
PRT: 1104 AA.
08XON7 PRELIMINARY;
01-MAR-2002 (Tremblrel, 20, Created)
01-MAR-2002 (Tremblrel, 20, Last sequence update)
01-JUN-2002 (Tremblrel, 21, Last annotation update)
DE Probable DNA-directed DNA polymerase III.
GN B10H4.020.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariales; Sordariaceae; Neurospora.
OX NCBI_Taxid=5141;
[1]
RP SEQUENCE FROM N.A.
RA Schultze U., Aign V., Hohenseel J., Brandt P., Fattmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RL EMBL: AL670010; CAD21389.1; -.
DR InterPro: IPR002064; DNA_pol_B.
DR InterPro: IPR004578; Pol2.
DR Pfam: PF03104; DNA_pol_B_exo; 1.
DR PRINTS: PR00106; DNAPOB.

```

DR SMART, SW00486; POLBc: 1.
 DR TIGRFAMs; TIGR00592; pol2: 1.
 DR PROSITE; PS00116; DNA_POLYMERASE_B; UNKNOWN_1.
 DR Dna-directed DNA polymerase.
 SQ SEQUENCE 1104 AA; 125507 MW; 3B5A1A0BFD26042F CRC64;

Query Match 17.6%; Score 737.5; DB 3; Length 1104;
 Best Local Similarity 28.6%; Pred. No. 3.1e-40;
 Matches 261; Conservative 154; Mismatches 311; Indels 185; Gaps 39;

17 GKEPQVYIINGIAENGERVYLIDRSFRPFY--ALLAPGAD--PKVAGRIALSRP---- 68
 123 GGGATTKLFGVTENGNSVLLHVKKFHYLVQAIVSFGDDCPKRALETQLMPTPAI 182
 69 -----KSPVIGVEDDKKRYGRRRLVLRTRVLPEAVREYRELKRV 110
 183 HSYVLTRENNYGFQGNQNPYIKVTVNDPKFLPRVRLIETNKA-----NMKGWKS- 235
 111 DGVEDVLEADIRFAMRYLIDHLPFTWYVEAE-----PLENGKGRVDKRY 158
 236 DGSIMTFD-DIGYLRFWDCSIGMSVNEAPAGSYQVINDGHSNCOLEAMSYR-----D 290
 159 LYSRPEPLXGEMLAPTKLPDLRIAPDIEVYSKQ--SPREPRDYVIA--VKTDDGDEV 216
 291 LIAHRP---VGEW---SKMAPRLISFDIECAQKGIFFEPQHDVIOIANIVTKYGEKK 344
 217 LFLAE--GKDRKPI-----REFVEYKRYDPDIIIVGNNHNPWYLLR 259
 345 PFMNVFCLDTTSPVIAQOISFREDMDLREGNRLLEQVDDPIIGYNIANPDPFYLID 404
 260 RAALIGIK-LDV-----TRVGAEPSTVSHGVSPGLNVDLYDAE 302
 405 RAKLTKVDFDFMSRTVRVSVAKETNFSKQMGNRDTKATNTN---GRLODMLQVQK 460
 303 MPEIKISLEEV-AEVLGVWKS--ERYIINMWEIPDYDDPKRPLLYQARDVARY 359
 461 DYOLRSYTLNSVCSHFLGEOKEDVHSHITELFE-----GTPESRRRLALYCLMDAVLP 515
 360 GLAEKI--LPAIOLSYVTGLPLDOVGMSVGRLEMYLIRAFKELVYN--RYEPRP 416
 516 RLMDKLSCLIENTEMARVATGVPFNFILARQOQKPLSOLFRLQKLVINLRPESSE 575
 417 TYGATVLEPLAGVHE-NIAYLDSSMYPMIMKYNVGDITVREGKCGEGCGEAPR- 474
 576 QYEGATVIEPTKGYIDVPLATIDPASLPSTIMQAHNLCTYTLI---KKRDIERMDLKD 631
 475 -----VGRPRRCPPGFKTVLERLLELRKRAEMKKYPPDSPRYLLDE 520
 632 EDYIVTPNEDMFVTTKOR-----KGLLAQILLEILSARQARREL-AVERKDPFKKAVLNG 685
 521 ROKALKVLNANASYGVWMSGARWYCRCAKAVTAMGRHLI-RTAINIRAKLGL----- 572
 686 ROLALKVANSYVGLGATNGTLCLEIASVTSFORQMIERTKKEVEERYTIANGYSHD 745
 573 -KVLVGPDSLFYTDPEKVB-----NFKIITKEELGEILKELVGRLEF 616
 746 AQLVGPDSVMWKGTGDLAEAMKLGODASEVYSKFIKPIK-----LEFEKVPYPI 799
 617 FTBAKKRYAGL-----EDGRIDVGEAVRGDMGLAEQGTQVLEVLKTSVYKAYE 671
 800 LIN-KKRYAGLWTKREKIDMDIKGIEYARDNCLLVQVLEKLRNMLIDQVPGQAQ 858
 672 YRKLYKELEBEGKVPLEKLVYKTLKRLBEYTEAPHVAAKRM--LSAGYRVSPEGKI 729
 859 YVKDTIADLLQNKIDMSKLVITKALK--ENYDAKAQAHVLLQNRKKKADASAGLDNR 916
 730 GYVIVGK--GGRISORAW-PFMVYDPSQIDVTVYVHOIIPALR-----ILGYGITEK 782
 917 AYVWKGATGSKRFRSEBDPIYVLEHNVPIIDTKYVLDNOLAPLRIFEPILG-----ET 971
 783 KTKASATGOKT 793
 972 KAKSLIHGDHT 982

RESULT 14
 ID OSQPS PRELIMINARY; PRT; 974 AA.
 AC OSQPS
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE DNA polymerase delta catalytic (large) chain.
 GN EC009_0430.
 OS Encephalitozoon cuniculi.
 OC Eukaryota; Microsporidia; Unikaryonidae; Encephalitozoon.
 NCBI_Taxid=6035,
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GB-M1,
 RA Genoscope,
 RL Submitted (APR-2001) to the EMBL/Genbank/DBJ databases.
 RV [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GB-M1,
 RA MEDLINE=21576510; PubMed=11719806;
 RA Kadinka M.D., Duprat S., Cornilliot E., Metenier G., Thomarat F.,
 RA Penster G., Barde V., Peyretilade E., Brotier P., Winkler P.,
 RA Delbac F., El Alaoui H., Peyret P., Saurin W., Gouy M.,
 RA Weissenbach J., Vivares C.P.
 RT "Genome sequence and gene compaction of the eukaryote parasite
 RT Encephalitozoon cuniculi.".
 RL Nature 414:450-453(2001).
 DR EMBL; AL590451, CAD27015.1,
 SQ SEQUENCE 974 AA; 111628 MW; 957E519E6EBB455 CRC64;

Query Match 17.5%; Score 732.5; DB 5; Length 974;
 Best Local Similarity 28.0%; Pred. No. 5.6e-40;
 Matches 249; Conservative 152; Mismatches 320; Indels 167; Gaps 33;

2 TEVYVTLDS--SYEVNGKEPQVYIINGIAENGERVYLIDRSFRPFVYLLAPG-----AD 54
 10 THIFVNDGSLGSLSC---PVFSIFONTYSGPFAVLMKNNFFYFVYBSNGKEKEED 65
 55 PKQVAGRIALSRPSPPIIGVEDDKRK---YGRPRVLRIRTVLPEAVREYRELKRV 110
 66 IKESVORLDV---KATLIEVAVMKQSLIGYEGKTRVYRLTNTPHVSTALKVLLSEG 121
 111 DGVED-----VLEADIRFAMRYLIDHLPFTWYVEAEPLNKGKGRVDKRYLVKSRP 164
 122 ISVGEKVRPRVYESNPFVLRFMCDLGIWMSYLVNRYEILDDHLTVASGY-ESLES 180
 165 EPLVGEALAPTKLPDLRIAPDIE-VYSKQSPREPRDYVIAVKTDDGDEVLFIAEGK 223
 181 LPLEGGYVV---LPLKVLISIDIEVAGNGPSSKCDPIIGI-----GNTLMLGDS 231
 224 DDRKPI-----REFVEYKRY---DPDIIIVGNNHNPWYLLR 259
 222 YASQIDFCLKETTGPVGNVWMEYETKELESKKKIPMELDPDVIAGVYKGDPIYILS 291
 260 RAALIGIK---LDVTRVGAEPSTVSHG-----VSPGRLANVDLYDAEMPEI 306
 292 RGEILGIBSFYVSGSKKAKTRDTWSSMFGSITTEVEIDGLIIMVNVIRDPFLK 351
 307 KIKLEEVAEYLVGMKKSERVIINMWEIPDYDDPKRPLLYQARDV---ATYGLAE 363
 352 RSYSLNVS--IHFLKEQKE-----DVFPSSIGELQSKNKRTRRIASVCLAD 397
 364 KILPFAI-----QLSYVTGLPLDOVGMSVGRLEMYLIRAFKELVYNRYE 412
 398 TVLEPLRLFNTLVNLYNTELSRVTGTIEYFFTRGAIKIFLTVYRAASKEDEFIIPD-ID 456
 413 --RPEETVYGAIVLEPLGVHEN--IAVLDPSMYPMIMKYNVGDITVREGKCGEGCG 469
 457 PFESNKTIFEGGVIEPRKGFYNKRVSMVDSISLYPSIMSHNL-----C 500

RESULT 15

O9HRV9 PRELIMINARY; PRT; 901 AA.

ID O9HRV9
AC O9HRV9;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE DNA polymerase B.
DN POLB1 OR VNGO521G.
OS Halobacterium sp. (strain NRC-1).
OC Archaeae; Euryarchaeota; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950.
RA Ng W.V., Kennedy S.P., Manairas G.G., Bergquist B., Pan M.,
RA Shukla H.D., Laasy S.R., Bailig N.S., Thorsson V., Shroghna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welis R., Goo Y.A.,
RA Leitner K., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angvine C.M., Dale H.,
RA Isehaberger T.A., Beck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
Alam H., Freltas T., Hou S., Daniels C.U., Dennis P.P., Omer A.D.,
Ebnhardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
"Genome sequence of Halobacterium species NRC-1,"
Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000)
RL -1- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE = N
CC CC
CC PHOSPHOPHATE + DNA(N).
CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
CC EMBL; AE005004; AAG19049.1; -;
DR InterPro; IPRO02064; dna_pol_B.
DR Pfam; PF00136; dna_pol_B_2.
DR Pfam; PF0104; dna_pol_B_exo; 1.
DR PRINTS; PR00106; DNAPOLB.
DR SMART; SM00486; POLBc; 1.
DR PROSITE; PS00116; DNA_POLYMERASE_B; 1.
KW DNA replication; DNA-binding; DNA-directed DNA polymerase;
KW Complete proteome.
SQ SEQUENCE 901 AA; 101987 MW; 213470487AB5E49D CRC64;

QY	61	RIRLNRKPSIIGVEDDK-----KXFG-PRVULIRIVLREAVREYELKAND	111
Db	113	-----DLTDVITTEGEGYESIRGEALTUKTIGTRPRVGNIRAF-----	1522
QY	112	GVEDVLREADIRFARVYLIDHLPFTYRYAEABELEKMGFRVYDKVLYVSRPEPLYGEA	171
Db	153	---DHYEADILFPRRLIIDMDI--TDGVRPAR-----RADDGALVYHNHEIAACEV	199
QY	172	LAPTKLPDLIRLADIEVYSKQSPREPERDVIYI-----	206
Db	200	AA-----DLRVNTPDIEVNRNGFPEDEGEPRVCLTSHSDYRDEYVAMLYAAPDATVADP	254
QY	207	-AVK-----TDDG---EVLFLAEGKODRPIREFVUEVYKRYDRIIVGNHNNHFDMPRYL	257
Db	255	TAVGYPDLTDADIDVRVPDETSAMHD-----AFSLYIEDTNPVYLTGNFDFPAPYL	309
QY	258	LRRARIIIGIK-----LDVTR--RVGAEPPTSVHGHVSVPGRILAND-LYDVAAE-EMPEIKIK	309
Db	310	IDRIDELDPRIHDHLDSDRLSRV-SEVWTSGMGSPNKGAVVDLXAYORTKXSELDST	368
QX	310	SLEBVAAY-IGVKKSSRRVLIINMWELPDYVDDPKGRPLLQYABDVRYTGLAEK--IL	366
Db	369	RLDVAGEOELGVGK--BRYP---GDIGDLWEDPPER--LLEYMLRDVELCVELDRKQSIY	421
QY	367	PEAIQLSVYVGLPELDQYGAMSVGRLEMYLIRAFKKKELVPMNRVERPEETYGALYLER	426
Db	422	AFMDARKVLQCKLED--ATTGVAIVMYVJLHAKFG-NFVLPSSGQODAEFEFGAATDP	478
QY	427	LRGVHENIAYLVDFFSSMYPNIIKTNVOPDPLVLRPGEKSGCGCEAEPVGRFRRCPCPF	486
Db	479	ITGVRENVSVLDKSLYFPMNAVITINASPETIKVDBEEBDDG--TYRTPGVH-FRKEBDGI	535
QY	487	FKTVLERLLELRKSVRAEMKKYPPDSEPYRLDEROKALVULANASYGYMGSGARYCR	546
Db	536	IREMWDELITREBEKKARRDDBHEDSDERYDRQCAAVKVINMSLYGVGMDFRLYDR	595
QY	547	EOAKVATMAGHGLRKTINILNARKLGLVUYIGDDLSLV-----TYDEKVE	592
Db	596	AMSGAVSTINEVVIDLFTGEQAAEBRGVVAAGDDTSVNLBLEDDNTKAELIAESFDIEDHI	655
QY	593	N--FIKIIKEEL---GFEIKLEKVKYKRLFTEAKKRYAG--LLEBGR-----LDIVGFE	639
Db	656	NAAYDEPARBEDLHADDHRFOJEFEK/LRRFFQAGKKRYAGHLYMKGKQVDDVDIDGFE	715
QY	640	AVRGDMCLAEVOTKQVEIYLYKTSBYNKAVEYRYKIVKELEBQKPIEKLVIWKTLSKR	699
Db	716	YORSIDIAPIETRVOKEVIDLVRGVDVAVEBYHGEIEFQEGDADLDDIGIPGIGKR	775
QY	700	LEEYTEAPHVAAK 714	
Db	776	LDNFTDPAOVGAK 790	

Search completed: April 23, 2003, 07:31:13
Job time : 47 secs

Query Match	17.2%;	Score 718.5;	DB 17;	Length 901;
Best Local Similarity	30.1%;	Pred. No. 4.2e-39;		
Matches 239;	Conservative 118;	Mismatches 293;	Indels 145;	Gaps 31

```
QY      3 EVVFVTLDDSSIEYVG--KEPQVIWGIAENGERVVLIIDRSRPFFAALLAPGADPKQAQ 60
          ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::
DB     58 ELMTQVDYTTEGGDRERPVLHFGRTTDEAEAHNRVHGRRPYFAFAPTALISED----- 112
```

Search completed: April 23, 2003, 07:31:13
Job time : 47 secs



[illegible]

DS		Claim 1; Fig 2; 72pp; English.
PT	DNA synthesis	- useful for
XX	DNA polymerases from extremely thermophilic bacteria	
XX	Pt	
DR	N-PSDB; AAX27283.	
XX	WPI; 1999-180490/15.	
XX	DR	
XX	Callen W, Matchur EJ;	
XX	(DIVE-) DIVERSA INC.	
XX	PA	
XX	PR	
XX	06-AUG-1997; 97US-0907166.	
XX	PF	
XX	06-AUG-1998; 98WO-US17152.	
XX	PD	
XX	18-FEB-1999.	
XX	MO9907837-A1.	
XX	OS	
XX	Pyrolobus fumarius.	
XX	KM	
XX	DNA polymerase; thermophilic bacteria; DNA synthesis.	
XX	DE	
XX	P. fumarius DNA polymerase 1PZ protein sequence.	
XX	DT	
AC	02-JUN-1999 (first entry)	
AA	AAV00936;	
ID	AAV00936 standard; Protein; 803 AA.	
RESULT 1		
AAV00936		

XX This sequence is a DNA polymerase of the invention, that was
 CC isolated from a thermophilic bacteria. The polymerases are used in DNA
 CC synthesis and as immunogens to raise antibodies (useful for affinity
 CC purification and to screen for related enzymes). Fragments of the DNA
 CC full-length sequences are used as probes to isolate related or
 CC polymerases catalyze DNA synthesis by the addition of deoxynucleotides to
 CC the 3' end of a polynucleotide chain, using a complementary
 CC polynucleotide strand as a template. The polymerases have optimum
 CC activity at over 60 deg. C and can renature and regain activity after
 CC exposure to temperatures above 70 deg. C.

XX Sequence 803 AA;

Query Match 100.0%; Score 4188; DB 20; Length 803;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 803; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTEVFTVLDSSYEYVGEKEPOVIYIWIAGNGBRVVLIIDRSFRPYFALLAPGADPKQVQAQ 60
 1 MTEVFTVLDSSYEYVGEKEPOVIYIWIAGNGBRVVLIIDRSFRPYFALLAPGADPKQVQAQ 60
 QY 61 RIRALSRRPKSPIIGVEDDKRRYFGRRPRVLRIRTVLPEAVREYRELKVNVDGVEVLEAD 120
 61 RIRALSRRPKSPIIGVEDDKRRYFGRRPRVLRIRTVLPEAVREYRELKVNVDGVEVLEAD 120
 DB 61 RIRALSRRPKSPIIGVEDDKRRYFGRRPRVLRIRTVLPEAVREYRELKVNVDGVEVLEAD 120
 QY 121 IRFAMRYLIDHDLPFTMYRVEAEPLNKGFRVDKXYLYKSPPEPLYGEALAPTKLPDL 180
 121 IRFAMRYLIDHDLPFTMYRVEAEPLNKGFRVDKXYLYKSPPEPLYGEALAPTKLPDL 180
 DB 121 IRFAMRYLIDHDLPFTMYRVEAEPLNKGFRVDKXYLYKSPPEPLYGEALAPTKLPDL 180
 QY 181 RILAFDIEVYSKQSSPRERDPYIVAVKTDGDEVLFIAEGKODRKPIREFVEYVKRYD 240
 181 RILAFDIEVYSKQSSPRERDPYIVAVKTDGDEVLFIAEGKODRKPIREFVEYVKRYD 240
 DB 181 RILAFDIEVYSKQSSPRERDPYIVAVKTDGDEVLFIAEGKODRKPIREFVEYVKRYD 240
 QY 241 PDIIYGYNNHFDMPYLLRRARILGKLDVTRVGAEPPTSVAHGVSVEGRNLVDLYDA 300
 241 PDIIYGYNNHFDMPYLLRRARILGKLDVTRVGAEPPTSVAHGVSVEGRNLVDLYDA 300
 DB 241 PDIIYGYNNHFDMPYLLRRARILGKLDVTRVGAEPPTSVAHGVSVEGRNLVDLYDA 300
 QY 301 EEMPEIKIKSLSEVAEYLVGMKKSSRVIIINWMEIPDYWDPKRPPILLQYAADVRATYG 360
 301 EEMPEIKIKSLSEVAEYLVGMKKSSRVIIINWMEIPDYWDPKRPPILLQYAADVRATYG 360
 DB 301 EEMPEIKIKSLSEVAEYLVGMKKSSRVIIINWMEIPDYWDPKRPPILLQYAADVRATYG 360
 QY 361 LAEKILPFAIQLSYVTGLPDDOYGAHSVGFLEWYLIRAAFKKKEIVPRKVRPEETRYG 420
 361 LAEKILPFAIQLSYVTGLPDDOYGAHSVGFLEWYLIRAAFKKKEIVPRKVRPEETRYG 420
 DB 361 LAEKILPFAIQLSYVTGLPDDOYGAHSVGFLEWYLIRAAFKKKEIVPRKVRPEETRYG 420
 QY 421 AIVEPLRAGHENIIVLDFSSMYPNIMIKVNVGPDITVIRGEGCGCGCEAPVYKHF 480
 421 AIVEPLRAGHENIIVLDFSSMYPNIMIKVNVGPDITVIRGEGCGCGCEAPVYKHF 480
 DB 421 AIVEPLRAGHENIIVLDFSSMYPNIMIKVNVGPDITVIRGEGCGCGCEAPVYKHF 480
 QY 481 RCPGPFKTYLERLLSLRKRYRAEKKYPPSPPEYRLIDRQALAVLANASYGWSG 540
 481 RCPGPFKTYLERLLSLRKRYRAEKKYPPSPPEYRLIDRQALAVLANASYGWSG 540
 DB 481 RCPGPFKTYLERLLSLRKRYRAEKKYPPSPPEYRLIDRQALAVLANASYGWSG 540
 QY 541 ARMYGRCACAVTAMGRHLIRTAIINARKLGKVIYSDTDSLFTVYDPEKVENFKI 600
 541 ARMYGRCACAVTAMGRHLIRTAIINARKLGKVIYSDTDSLFTVYDPEKVENFKI 600
 DB 541 ARMYGRCACAVTAMGRHLIRTAIINARKLGKVIYSDTDSLFTVYDPEKVENFKI 600
 QY 601 ELGFEIKLEKYKRLPFEAKRGYAGLLEDSRIDVFEAVRDWCELAKEVQTVAVEIV 660
 601 ELGFEIKLEKYKRLPFEAKRGYAGLLEDSRIDVFEAVRDWCELAKEVQTVAVEIV 660
 DB 601 ELGFEIKLEKYKRLPFEAKRGYAGLLEDSRIDVFEAVRDWCELAKEVQTVAVEIV 660
 QY 661 LKTSVNVKAVEYRKIVKELSEGKPIEKLVIWMTLSKRLSEETTEPPIVVAKKMLAG 720
 661 LKTSVNVKAVEYRKIVKELSEGKPIEKLVIWMTLSKRLSEETTEPPIVVAKKMLAG 720
 DB 661 LKTSVNVKAVEYRKIVKELSEGKPIEKLVIWMTLSKRLSEETTEPPIVVAKKMLAG 720
 QY 721 YRSPGPKIGYIVYKGGGRIISORAMPYMWDPDSQIDVTYVHOHIIIPALNIIIGYBIT 780
 721 YRSPGPKIGYIVYKGGGRIISORAMPYMWDPDSQIDVTYVHOHIIIPALNIIIGYBIT 780
 DB 721 YRSPGPKIGYIVYKGGGRIISORAMPYMWDPDSQIDVTYVHOHIIIPALNIIIGYBIT 780
 QY 781 EKKLKASATGOKTLFPLAKSK 803
 781 EKKLKASATGOKTLFPLAKSK 803

DB 781 EKKLKASATGOKTLFPLAKSK 803

RESULT 2

AAE22118 ID, AAE22118 standard; Protein; 803 AA.

AAE22118;

25-JUL-2002 (first entry)

Pyrobolus fumaria DNA polymerase, 1PY2.

DNA polymerase; thermostable; enzyme.

Pyrobolus fumaria.

MO200220735-A2.

14-MAR-2002.

06-SEP-2001; 2001MO-US28007.

06-SEP-2000; 2000US-0656309.

(DIYE-) DIVERSA CORP.

Callen W, Mathur EJ, Short JM,

WPI; 2002-362247/39.

N-PDB; AAD35187.

New thermostable polymerase useful for sequencing DNA, amplifying

double stranded DNA, or incorporating a non-natural nucleotide or a

nucleotide analog into a DNA molecule

Claim 42; Fig 1A-E; 16tp; English.

The invention relates to thermostable DNA polymerases having high

temperature polymerase activity, such as those derived from

Pyrobolus fumaria and nucleic acid molecules encoding such polymerases.

Polymerases are useful for catalyzing the formation or repair of a

nucleic acid sequence and for modifying small molecules. They are also

useful for sequencing DNA molecules, for preparing cDNA from mRNA, for

amplifying double stranded DNA molecules and for incorporating non-

natural nucleotides or nucleotide analogues into a DNA molecule. The

present sequence is Pyrobolus fumaria DNA polymerase.

Sequence 803 AA;

Query Match 100.0%; Score 4188; DB 23; Length 803;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 803; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTEVFTVLDSSYEYVGEKEPOVIYIWIAGNGBRVVLIIDRSFRPYFALLAPGADPKQVQAQ 60
 1 MTEVFTVLDSSYEYVGEKEPOVIYIWIAGNGBRVVLIIDRSFRPYFALLAPGADPKQVQAQ 60
 DB 1 MTEVFTVLDSSYEYVGEKEPOVIYIWIAGNGBRVVLIIDRSFRPYFALLAPGADPKQVQAQ 60
 QY 61 RIRALSRRPKSPIIGVEDDKRRYFGRRPRVLRIRTVLPEAVREYRELKVNVDGVEVLEAD 120
 61 RIRALSRRPKSPIIGVEDDKRRYFGRRPRVLRIRTVLPEAVREYRELKVNVDGVEVLEAD 120
 DB 61 RIRALSRRPKSPIIGVEDDKRRYFGRRPRVLRIRTVLPEAVREYRELKVNVDGVEVLEAD 120
 QY 121 IRFAMRYLIDHDLPFTMYRVEAEPLNKGFRVDKXYLYKSPPEPLYGEALAPTKLPDL 180
 121 IRFAMRYLIDHDLPFTMYRVEAEPLNKGFRVDKXYLYKSPPEPLYGEALAPTKLPDL 180
 DB 121 IRFAMRYLIDHDLPFTMYRVEAEPLNKGFRVDKXYLYKSPPEPLYGEALAPTKLPDL 180
 QY 181 RILAFDIEVYSKQSSPRERDPYIVAVKTDGDEVLFIAEGKODRKPIREFVEYVKRYD 240
 181 RILAFDIEVYSKQSSPRERDPYIVAVKTDGDEVLFIAEGKODRKPIREFVEYVKRYD 240
 DB 181 RILAFDIEVYSKQSSPRERDPYIVAVKTDGDEVLFIAEGKODRKPIREFVEYVKRYD 240
 QY 241 PDIIYGYNNHFDMPYLLRRARILGKLDVTRVGAEPPTSVAHGVSVEGRNLVDLYDA 300
 241 PDIIYGYNNHFDMPYLLRRARILGKLDVTRVGAEPPTSVAHGVSVEGRNLVDLYDA 300
 DB 241 PDIIYGYNNHFDMPYLLRRARILGKLDVTRVGAEPPTSVAHGVSVEGRNLVDLYDA 300

```

QY 301 BEMPEIKIKSLEVAEYLVGKKSSERVIIIMWEIPDYWDDBPKRPLLIQVARDVDRATY 360
|||
DB 301 BEMPEIKIKSLEVAEYLVGKKSSERVIIIMWEIPDYWDDBPKRPLLIQVARDVDRATY 360
|||
QY 361 LAEKILPFAIQLSVYVGLPLDQVGMSSVGRLEWYLIRAAFKMKELVNNRVERPEETRYG 420
|||
DB 361 LAEKILPFAIQLSVYVGLPLDQVGMSSVGRLEWYLIRAAFKMKELVNNRVERPEETRYG 420
|||
QY 421 AIVLEPLRGVHENIAVLDSSMYPNIMIKYVGPDTLVPRGEKCGECGCWEAPEVKRFR 480
|||
DB 421 AIVLEPLRGVHENIAVLDSSMYPNIMIKYVGPDTLVPRGEKCGECGCWEAPEVKRFR 480
|||
QY 481 RCPGPFKTYLERLLELRKRVRAEMKKYPPDSPERYLLDEROKALKYLANASYGMGMSG 540
|||
DB 481 RCPGPFKTYLERLLELRKRVRAEMKKYPPDSPERYLLDEROKALKYLANASYGMGMSG 540
|||
QY 541 ARMYCRECAKAVTAMGRHLIRTAIINARKGLKVIYGDTSLFVYDPEKVENFIKIIE 600
|||
DB 541 ARMYCRECAKAVTAMGRHLIRTAIINARKGLKVIYGDTSLFVYDPEKVENFIKIIE 600
|||
QY 601 ELGFEIKLEKYKRLFFTEAKKRYAGLLEDGRIDIVGFEAVRGDMCELAKEVQTRVEIV 660
|||
DB 601 ELGFEIKLEKYKRLFFTEAKKRYAGLLEDGRIDIVGFEAVRGDMCELAKEVQTRVEIV 660
|||
QY 661 LKTSVNKAVEYVRKIYKELEEGKPIEKLVIWKTLSKRLSEYTTTEAPHVAAKRMISAG 720
|||
DB 661 LKTSVNKAVEYVRKIYKELEEGKPIEKLVIWKTLSKRLSEYTTTEAPHVAAKRMISAG 720
|||
QY 721 YRVSFGDKIGVIVKGGGRISQRAFPYFVKDPSQIDVTVYVDHQLIPALRLIGYFGIT 780
|||
DB 721 YRVSFGDKIGVIVKGGGRISQRAFPYFVKDPSQIDVTVYVDHQLIPALRLIGYFGIT 780
|||
QY 781 EKKIKASATGQKTLFDPLAKSK 803
|||
DB 781 EKKIKASATGQKTLFDPLAKSK 803
|||

RESULT 3
ID AAE22119 standard; Protein; 803 AA.
XX AC AAE22119;
XX DT 25-JUL-2002 (first entry)
XX Pyrolobus fumaria DNA polymerase, PLF1831.

DNA polymerase; thermostable; enzyme.
XX Pyrolobus fumaria.
XX WO200220735-A2.
XX PD 14-MAR-2002.
XX PF 06-SEP-2001; 2001WO-US28007.
XX PR 06-SEP-2000; 2000US-0656309.
XX PA (DIVE-) DIVERSA CORP.
XX PI Callen W, Mathur EJ, Short JM;
XX WPI; 2002-362247/39.
XX DR N-PSDB; AAD35188.
XX PT New thermostable polymerase useful for sequencing DNA, amplifying
XX double stranded DNA, or incorporating a non-natural nucleotide or a
XX nucleotide analog into a DNA molecule
XX PS Claim 42; Fig 11; 161p; English.
XX

```

CC The invention relates to thermostable DNA polymerases having high
 CC temperature polymerase activity, such as those derived from
 CC Pyrolobus fumaria and nucleic acid molecules encoding such polymerases.
 CC Pyrolobus are useful for catalysing the formation or repair of a
 CC nucleic acid sequence and for modifying small molecules. They are also
 CC useful for sequencing DNA molecules, for preparing cDNA from mRNA, for
 CC amplifying double stranded DNA molecules and for incorporating non-
 CC natural nucleotides or nucleotide analogues into a DNA molecule. The
 CC present sequence is Pyrolobus fumaria DNA polymerase.

XX Sequence 803 AA;

Query Match 99.4%; Score 4164; DB 23; Length 803;

Best Local Similarity 99.3%; Pred. No. 0; Mismatches 1; Indels 0; Gaps 0;

Matches 797; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 MTEVFTVLDSSYEYVKEPQVITWGIABGSEVVLIDRSFRFPYFALLAPGADPKQVAQ 60
|||
DB 1 MTEVFTVLDSSYEYVKEPQVITWGIABGSEVVLIDRSFRFPYFALLAPGADPKQVAQ 60
|||
QY 61 RIRALSRPKSPIIGVEDDKKRYGPRPRVLRIRTVLPEAVREYRELKONVGVEDVLEAD 120
|||
DB 61 RIRALSRPKSPIIGVEDDKKRYGPRPRVLRIRTVLPEAVREYRELKONVGVEDVLEAD 120
|||
QY 121 IRFAMRYLIDHDLPPFTWYVVEAEPLENNKGFVVDKVIYKSRPEPLYGALAPTKLPDL 180
|||
DB 121 IRFAMRYLIDHDLPPFTWYVVEAEPLENNKGFVVDKVIYKSRPEPLYGALAPTKLPDL 180
|||
QY 181 RILAFDIEVSKSGSPRPERDPVIVIAVKTDDGDEVLFINEGDDKRPILFEFVEYVRKD 240
|||
DB 181 RILAFDIEVSKSGSPRPERDPVIVIAVKTDDGDEVLFINEGDDKRPILFEFVEYVRKD 240
|||
QY 241 PDIIIVGNNHFPWYLLRPARILGILKLDVTRVGAEPPTSVMGHVSVPRGLNVDLYDA 300
|||
DB 241 PDIIIVGNNHFPWYLLRPARILGILKLDVTRVGAEPPTSVMGHVSVPRGLNVDLYDA 300
|||
QY 301 BEMPEIKIKSLEVAEYLVGKKSSERVIIIMWEIPDYWDDBPKRPLLIQVARDVDRATY 360
|||
DB 301 BEMPEIKIKSLEVAEYLVGKKSSERVIIIMWEIPDYWDDBPKRPLLIQVARDVDRATY 360
|||
QY 361 LAEKILPFAIQLSVYVGLPLDQVGMSSVGRLEWYLIRAAFKMKELVNNRVERPEETRYG 420
|||
DB 361 LAEKILPFAIQLSVYVGLPLDQVGMSSVGRLEWYLIRAAFKMKELVNNRVERPEETRYG 420
|||
QY 421 AIVLEPLRGVHENIAVLDSSMYPNIMIKYVGPDTLVPRGEKCGECGCWEAPEVKRFR 480
|||
DB 421 AIVLEPLRGVHENIAVLDSSMYPNIMIKYVGPDTLVPRGEKCGECGCWEAPEVKRFR 480
|||
QY 481 RCPGPFKTYLERLLELRKRVRAEMKKYPPDSPERYLLDEROKALKYLANASYGMGMSG 540
|||
DB 481 RCPGPFKTYLERLLELRKRVRAEMKKYPPDSPERYLLDEROKALKYLANASYGMGMSG 540
|||
QY 541 ARMYCRECAKAVTAMGRHLIRTAIINARKGLKVIYGDTSLFVYDPEKVENFIKIIE 600
|||
DB 541 ARMYCRECAKAVTAMGRHLIRTAIINARKGLKVIYGDTSLFVYDPEKVENFIKIIE 600
|||
QY 601 ELGFEIKLEKYKRLFFTEAKKRYAGLLEDGRIDIVGFEAVRGDMCELAKEVQTRVEIV 660
|||
DB 601 ELGFEIKLEKYKRLFFTEAKKRYAGLLEDGRIDIVGFEAVRGDMCELAKEVQTRVEIV 660
|||
QY 661 LKTSVNKAVEYVRKIYKELEEGKPIEKLVIWKTLSKRLSEYTTTEAPHVAAKRMISAG 720
|||
DB 661 LKTSVNKAVEYVRKIYKELEEGKPIEKLVIWKTLSKRLSEYTTTEAPHVAAKRMISAG 720
|||
QY 721 YRVSFGDKIGVIVKGGGRISQRAFPYFVKDPSQIDVTVYVDHQLIPALRLIGYFGIT 780
|||
DB 721 YRVSFGDKIGVIVKGGGRISQRAFPYFVKDPSQIDVTVYVDHQLIPALRLIGYFGIT 780
|||
QY 781 EKKIKASATGQKTLFDPLAKSK 803
|||
DB 781 EKKIKASATGQKTLFDPLAKSK 803
|||

```

```

RESULT 4
AAE22120
ID AAE22120 standard; Protein; 799 AA.
XX
AC AAE22120;
XX
DT 25-JUL-2002 (first entry)
XX
DE Pyrolobus fumaria DNA polymerase consensus sequence.
XX
KW DNA polymerase; thermostable; enzyme.
XX
OS Pyrolobus fumaria.
XX
BN W0200220735-A2.
XX
PD 14-MAR-2002.
XX
PF 06-SEP-2001; 2001MO-US28007.
XX
PP 06-SEP-2000; 2000US-0656309.

(DIVE-) DIVERSA CORP.
XX
PI Callen W, Mathur EJ, Short JM;
XX
DR WPI; 2002-362247/39.
XX
PT New thermostable polymerase useful for sequencing DNA, amplifying
PT double stranded DNA, or incorporating a non-natural nucleotide or a
PT nucleotide analog into a DNA molecule.
XX
PS Disclosure; Fig 11; 16pp; English.
XX
CC The invention relates to thermostable DNA polymerases having high
CC temperature polymerase activity, such as those derived from
CC Pyrolobus fumaria and nucleic acid molecules encoding such polymerases.
CC Polymerases are useful for catalyzing the formation or repair of a
CC nucleic acid sequence and for modifying small molecules. They are also
CC useful for sequencing DNA molecules, for preparing cDNA from mRNA, for
CC amplifying double stranded DNA molecules and for incorporating non-
CC natural nucleotides or nucleotide analogues into a DNA molecule. The
CC present sequence is Pyrolobus fumaria DNA polymerase consensus sequence.
XX
SQ
Sequence 799 AA;
Query Match 98.4%; Score 4122; DB 23; Length 799;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 798; Conservative 1; Mismatches 0; Indels 4; Gaps 4;

1 MTEVVFVLBSSEVVGKEPOVIINGIENGGERVVLIDRSFRFPYFALLAPADPKVOAQ 60
DB 1 MTEVVFVLBSSEVVGKEPOVIINGIENGGERVVLIDRSFRFPYFALLAPADPKVOAQ 60
QY 61 RIRLSRPSKPIIGVEDDGRKIFGRPRVLAIRTVLEAVREIRELVKRVNGVEVELEAD 120
DB 61 RIRLSRPSKPIIGVEDDGRKIFGRPRVLAIRTVLEAVREIRELVKRVNGVEVELEAD 120
QY 121 IRFAMRYLIDHDLFFPTWYRVVAEPLKNGKFRVDKYYLVKSPPELYGEALAPTKLPDL 180
DB 121 IRFAMRYLIDHDLFFPTWYRVVAEPLKNGKFRVDKYYLVKSPPELYGEALAPTKLPDL 180
QY 181 RIILAFDIEVYSKQSPRPDRDVIIVAKTDDGDEVLFIAEGKDDKPIREFVEYVKRYD 240
DB 181 RIILAFDIEVYSKQSPRPDRDVIIVAKTDDGDEVLFIAEGKDDKPIREFVEYVKRYD 240
QY 241 PDIIIVGYNHNFDPMPYLIRRAILIGIKDVTRRVGAEPPTSVGHVSVGRLAVDLXYA 300
DB 241 PDIIIVGYNHNFDPMPYLIRRAILIGIKDVTRRVGAEPPTSVGHVSVGRLAVDLXYA 300
QY 301 EEMPEIKIKSLAEVAAYLGVMMKSSRVIIINMEIPDVMDDPKRPILLQYAADDPRAAYG 360
DB 301 EEMPEIKIKSLAEVAAYLGVMMKSSRVIIINMEIPDVMDDPKRPILLQYAADDPRAAYG 360

```

```

QY 361 LAEKILPPIAIOISVVTGLPIDOYGAMSVGRLEWYLIRAPFMKEIVPNRVERPEETYG 420
DB 361 LAEKILPPIAIOISVVTGLPIDOYGAMSVGRLEWYLIRAPFMKEIVPNRVERPEETYG 420
QY 421 AIVLEPLAGWHEINAVLDSSSMYPNIMIKYNGDPTVIRPCKGEGCGEAEAVYGRRR 480
DB 421 AIVLEPLAGWHEINAVLDSSSMYPNIMIKYNGDPTVIRPCKGEGCGEAEAVYGRRR 480
QY 481 RCPDGFETVIRLEILRKAVRAEMKTPDPSPEYRLDEROKALKYLANASYGIWMSG 540
DB 480 RCPDGFETVIRLEILRKAVRAEMKTPDPSPEYRLDEROKALKYLANASYGIWMSG 540
QY 541 ARMYCRECAKAVTAMGHHILRTAIIINIRKLGLKVIYGDTSLSFTYDPEKVENFIKIKE 600
DB 540 ARMYCRECAKAVTAMGHHILRTAIIINIRKLGLKVIYGDTSLSFTYDPEKVE-FIKII-E 596
QY 601 ELGFEIKLEKYKSLPFEAKKRYAGLLEDGRIDIVGFEAVRGDMCELAKEVOTKVEIV 660
DB 597 ELGFEIKLEKYKSLPFEAKKRYAGLLEDGRIDIVGFEAVRGDMCELAKEVOTKVEIV 656
QY 661 LKTSEVNKAVRYVARIYKELLEGKVPTEKLVIMKTLKRLSEYTTAPHYVAAKRLSAG 720
DB 657 LKTSVKNKAVRYVARIYKELLEGKVPTEKLVIMKTLKRLSEYTTAPHYVAAKRLSAG 716
QY 721 YVSPGDKIGYIVYKGGGRISORAMPYFMVNDPSQIDVTYVVDHQTIIIPALRIIGYFGIT 780
DB 717 YVSPGDKIGYIVYKGGGRISORAMPYFMVNDPSQIDVTYVVDHQTIIIPALRIIGYFGIT 776
QY 781 EKKLRASATGQKITLPDLAKSK 803
DB 777 EKKLRASATGQKITLPDLAKSK 799

RESULT 5
AA92523
ID AA92523 standard; Protein; 803 AA.
XX
AC AA92523;
XX
DT 12-JUL-1996 (first entry)
XX
DE Pyrodicticum occultum 803 amino acid DNA polymerase gene.
XX
KW DNA pol; base pair; thermally stable; exonuclease activity.
XX
OS Pyrodicticum occultum DSM2709.
XX
PN JP07327684-A.
XX
PD 19-DEC-1995.
XX
PF 09-JUN-1994; 94JP-0150591.
XX
PR 09-JUN-1994; 94JP-0150591.
XX
PA (TAKI ) TAKARA SHUZO CO LTD.
XX
DR WPI; 1996-072342/08.
XX
PT DNA encoding DNA polymerase - useful for prodn. of thermally stable
PT enzyme
XX
PS Claim 1; Page 11-13; 23pp; Japanese.
XX
CC AA92522 and AA92523 are 914 amino acid and 803 amino acid DNA
CC polymerases, respectively. They are derived from Pyrodicticum
CC occultum and are thermostable DNA polymerases. Also disclosed
CC are DNA polymerase genes which hybridize to the above genes.
XX
SQ
Sequence 803 AA;
Query Match 72.4%; Score 3033.5; DB 17; Length 803;

```

Best Local Similarity 70.8%; Pred. No. 1.6e-276;
Matches 571; Conservative 106; Mismatches 121; Indels 7; Gaps 5;

```

Qy 1 MTEVY-FTVLSSYEVGKEPOVIMIGIAENGERVYLIDRSFRPFYALLAPGADP--KQ 57
Db 1 MTEIIEFVLDDSSYEILKEPEVILMGITLDGRVLLDHRFRPFYALLARGEDWEE 60
Qy 58 VAORIRALSRPKSPIIGVEDDKRYKFRPRVLRIRTVLPEAVREYRELKXNVGVEDVL 117
Db 61 IASIRRLSVKSPILIDAKPLDKRYFGPRKAVKITMIPESVRHYRAVAKKIEGEDSL 120
Qy 118 EADIRFAMRYLIDHLEPFTYRVEAEPLKNGRVDKYLKYSRPEPLYGEALAPTKL 177
Db 121 EADIRFAMRYLIDKRLYPTFYRIPEVDAGRNPGRVDRVYKVGADPEPL--ADITRIDL 178
Qy 178 PDILIAFDIEVYSGKQSPREPDPIYIAVKTDDGDEVLIAGKDKDRKPIREFVEYK 237
Db 179 PPMRLVAFDIEVYSRSGSPNARDPVIIVSLRDSGKERLLIABGHDRRVLREFVEYR 238
Qy 238 RYDPDIIYGNNNHFDMPYLLRRARILGILDTYRVAEPFTTSVGHVSGRLNVLDLY 297
Db 239 AFDPDIIYGNNSHFDMPYLLMERARRLGILDTYRVAEPFTTSVGHVSGRLNVLDLY 298
Qy 298 DYAEEMPEIKIKLSEVAEYLGVNKKSERVIINWMEIPDYDDPKRPLLQYARDVRA 357
Db 299 DYAEEMPEIKIKLSEVAEYLGVNKKSERVILIMWRIPEYWDDEKQQLERVALDVRA 358
Qy 358 TYGLAEKILPFAIQLSTVYGLPLDQVGMVSGFRLEWYLIRAAFKKELVNNRVERPEET 417
Db 359 TYGLAEKMLPFAIQLSTVYGLPLDQVGMVSGFRLEWYLIRAAVDMNVLVNNRVERRGS 418
Qy 418 YRGAIVLEPLRGHENIAVLDPSSMYENIMIKYNNVGPDLVRPGEKGC-GCWEAPEVK 476
Db 419 YKGAIVLKLPGVHENVAVLDPSSMYEIMIKYNNVGPDLVDDSECKYGCYVAPEVG 478
Qy 477 HRFRCRPGPFKTVLERLLERKRVRAEMKKYPPDSEPYLLDEROKALKYLANASYGM 536
Db 479 HRFRRSPGPFKTVLERLLERKRVQKEKKEPPDSEPYLLYDEROKALKYLANASYGM 538
Qy 537 GWSGARWYCRCAKAVTAMGHLIRTAIINAKIGLKIYDGTDSLPTTYPEKYNFIK 596
Db 539 GWSGARWYCRCAKAVTAMGHLIRTAIIVARKIGLKIYDGTDSLPTTYPEKYNFIK 598
Qy 597 IIKELGFEILKEVYKRLFTFAKKRYAGLLEDRIDIVGEAVRGWCELAKEVOTKV 656
Db 599 FVEKELGEIIRDKIKYKVFTEAKKRYGLLEDRIDIVGEAVRGWCELAKEVDEKA 658
Qy 657 VEIYLTSEVNAKAVEYKVIKELBEGRVPIEKLIVKTLSEKLEBYTTEAPHVAAKRM 716
Db 659 AEIYLTNGVNDKALSYIREVILKOLREGKVPITKLIIMKTLSEKLEBYTTEAPHVAAARM 718
Qy 717 LSAAYRVSFGKIGVYVYKGGRIISQAMPYPMYKQDPSQIDVTYVHDHIIIPALRIIGY 776
Db 719 KEAYEVSPPGKGVYVYKGGSVSSRAVYPMYVDPSTIDVNYIHDQIIPALRIISY 777
Qy 777 FGITEKKLAKASATGOKTLFDPLAKK 801
Db 778 FGITEKKLAKAATVORSLEDFPASK 802

```

RESULT 6
AAR60883
ID AAR60883 standard; Protein: 803 AA.
XX
AC AAR60883;
XX
DT 19-JUL-1995 (first entry)
XX
DE P. occultum DNA polymerase.
XX
KM Polymerase chain reaction; PCR; amplify; primer; probe;
KM Pyrodictum occultum; DNA polymerase; P. abyssi; condensation;
KM DNA sequencing; nick translation; reverse transcription;

KM heat stability; heat resistant polymerase; DNA helix stabiliser.

XX Pyrodictum occultum.

XX EP624641-A.

PD 17-NOV-1994.

PF 02-MAY-1994; 94EP-0106811.

PR 14-MAY-1993; 93US-0062368.

PA (HOFF) HOFFMANN LA ROCHE & CO AG F.

PT Gelfand DH, Wang AM;

XX WPI; 1994-350781/44.

DR N-PDB; AAQ73844.

PT New heat stable nucleic acid polymerase from Pyrodictum species

PT - and relates DNA, vectors and transformed hosts, resistant to

PT denaturation at 100 deg. C, useful for nucleic acid amplification,

PT in DNA sequencing etc.

PS Claim 10, Page 33-35; 45pp; English.

CC This sequence represents the Pyrodictum occultum DNA polymerase gene.

CC The P. occultum DNA polymerase gene shows regions of homology to the

CC P. abyssi DNA polymerase genes, and the primers given in AAQ73824-25

CC are based on these regions. The amplified DNA encodes this DNA

CC polymerase which catalyses condensation of dNTPs to form complementary

CC strand DNA. This polymerase may be used in nucleic acid amplification,

CC partic. PCR, but also for DNA sequencing, nick translation and reverse

CC transcription. The polymerase has better heat stability than known

CC heat resistant polymerases. Temperatures of 100 deg. C can be

CC tolerated, facilitating strand separation and elimination of the

CC need for DNA helix stabilisers.

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

The present sequence is that of the novel DNAPolymerase of the anaerobic hyperthermophilic Archaeon, *Pyrobaculum islandicum* DSM24184. The sequence was deduced from a DNA polymerase gene (see AA50678) isolated from *P. islandicum* chromosomal DNA. The enzyme includes 6 conserved motifs, indicative of a family B DNA polymerase. A method for producing recombinant DNA polymerase using *Escherichia coli* BL21 cells transformed by a vector including the *P. islandicum* gene is claimed. The *P. islandicum* DNA polymerase is thermostable, having a half-life of 30-40 min at 100 degree C and over 10 hr at 90 degree C. It can be used for *in vitro* amplification of DNA fragments and for DNA sequencing (claimed). The enzyme also has 3'-5' proofreading exonuclease activity, providing a substantially lower base incorporation error rate than non-proofreading DNA polymerases.

```

QY 1DSSJEWGKEQOVIWIGJLAENGEEVVLIDSSFPYF---ALLAPGDPROVAQRIRAL 65
QY 9LDSJEWGKEQOVIWIGJLAENGEEVVLIDSSFPYF---ALLAPGDPROVAQRIRAL 65
Db 8LDITAVVGSVEPILFGILSGSERVVLIDSSFPYFYFDCAVCEPAA-----LKTAL 60
QY 66SHPKSPILIGVEDDKKRTGGRPRVRILRTVLPEAVREIRELVKNVDGVEDYLEADIRFAM 1255
Db 61SR-VAPIDPDQVJVERBFLGRSKKFLKIAKIPEDVRKILREAMSI PRVSGYVEDIDRFYM 119
QY 126RTLIDHDLPPFTWYRVVBAEPLTNKMGFPVDKY--YLVSRPEPLYG--EALAPKLPDLR 161
Db 120RIMIDMGVFPCCNNVAEVE-----GSLGGIPTYVSQ-----WYIDGEPBS---LK 166
QY 182ILAFDIEVYSKQGSBRPERDPYVIAVKTDDDEVLFLAEKDDRKPIREFVEVKKRYP 241
Db 167VNAFDIEVYNERGSDPDIRDPVMLAIKTNDQHEEVEFASGXDDRGVYRVPDIRSYDP 226
QY 242DIIYVNNNNHPWVPLRLARAILGLKLDVTRVGAEPRTTSHGVAVSPGRANDLHYAB 301
Db 227DVIYGNNSGFWPLRLVBEKAAVGPRLVD--RLSNPQDSYYGMSYVGRANDLVNIVE 265

```

RESULT 9
 AAY52022
 ID AAY52022 standard; Protein; 781 AA.
 AC AAY52022;
 XX
 DT 12-JUL-2000 (first entry)
 DE A. fulgidus AF0497 protein.
 XX
 KW Polymerase; thermostable; nucleic acid elongation; sliding clamp protein.
 KM amplification; reverse transcription.
 XX
 OS Archaeoglobus fulgidus.
 XX
 PN DE19840771-A1.
 PD 10-FEB-2000.
 XX
 PF 07-SEP-1998; 98DE-1040771.
 XX
 PR 06-AUG-1998; 98DE-1035653.
 XX
 PA (LION-) LION BIOSCIENCE AG.
 XX
 PI Voss H, Moeckel G, Kober I, Kilger C;
 DR WPI; 2000-207149/19.
 XX
 PT A thermostable in vitro polymerase complex for template-dependent
 PT elongation of nucleic acids in amplification or reverse transcription
 XX methods -
 XX
 PS Disclosure; Page 72-75; 152PP; German.
 XX
 CC This invention describes a novel thermostable in vitro complex for

RESULT 12

AAE22114

ID AAE22114

XX AAE22114: standard; Protein; 768 AA.

XX AAE22114:

XX AAE22114:

XX AAE22114:

XX AAE22114:

XX AAE22114:

XX AAE22114:

XX AAE22114:

XX AAE22114:

XX AAE22114:

XX AAE22114:

XX AAE22114:

XX AAE22114:

XX AAE22114:

XX AAE22114:

XX AAE22114:

XX AAE22114:

XX AAE22114:

XX AAE22114:

XX AAE22114:

XX AAE22114:

XX AAE22114:

XX AAE22114:

XX AAE22114:

XX AAE22114:

XX AAE22114:

XX AAE22114:

XX AAE22114:

XX AAE22114:

XX AAE22114:

XX AAE22114:

XX AAE22114:

XX AAE22114:

XX AAE22114:

XX AAE22114:

XX AAE22114:

XX AAE22114:

Archaeoglobus lithotrophicus DNA polymerase.

DNA polymerase; thermostable; enzyme.

Archaeoglobus lithotrophicus.

MO200220735-A2.

14-MAR-2002.

06-SEP-2001; 2001WO-US28007.

06-SEP-2000; 2000US-0656309.

(DIVE-) DIVERSA CORP.

Callen W, Mathur EJ, Short JM,

WPI; 2002-362247/39.

N-PSDB; AAD35183.

New thermostable polymerase useful for sequencing DNA, amplifying double stranded DNA, or incorporating a non-natural nucleotide or a nucleotide analog into a DNA molecule.

Claim 42; Page 139-141; 161pp; English.

The invention relates to thermostable DNA polymerases having high temperature polymerase activity, such as those derived from Pyrobolus fumaria and nucleic acid molecules encoding such polymerases. Polymerases are useful for catalysing the formation or repair of a nucleic acid sequence and for modifying small molecules. They are also useful for sequencing DNA molecules, for preparing cDNA from mRNA, for amplifying double stranded DNA molecules and for incorporating non-natural nucleotides or nucleotide analogues into a DNA molecule. The present sequence is Archaeoglobus lithotrophicus DNA polymerase.

Sequence 768 AA:

Query Match 37.3%; Score 1563; DB 23; Length 768;

Best Local Similarity 41.3%; Pred. No. 7.4e-138; Matches 338; Conservative 154; Mismatches 273; Indels 54; Gaps 18;

QY 1 MTEVFTVLDSSFEVVGKEPOVIWGIANGSERVVLIDRSPRYFALIAAPADPKOVAQ 60

DB 1 MTKKGLDADYITENDRAVIRLWCKDEG-IFIAVDSFOFPFALKEBGTAEIVK 59

QY 61 -RIPALSRPKSPIIGVEDDKRYFGPRPVLRITVLPEAVREYRELVXNVGVADVLEA 119

DB 60 IKVOTKKEVITP-LKVEETTAKNLGRVEVEFKIYARHCHVPKLRREVSO--YLEIREA 115

QY 120 DIRFAMRYLIDHLPFTTWYVAEPLNKGQFVADVLYLVK--SRPEPLYGEALAPKL 177

DB 116 DIPFAYRYLIDKMLACMDGVIEGVERREK-GLRCYKIKRIERDSHOD-----F 163

QY 178 PDLRLAFDIEVVSQKSGSPRPERDPIYAVATDGDGVLFIAGGDDRRKPIREFVEYVK 237

DB 164 PELKVMAFDCEMLSFVGMPDEPKEDPIYISIKSGYEILL--NGDNERELLTRRVKIR 220

QY 228 RYDDIIVGNNHHPMVLARARILGKIDVRRGAEPTSVGHVSGRANLTX 297

DB 221 DIDDIIIVGNOSDFMPPYIKKAEKLVKIDIGR--DRSELIRGGRPKIKGRANVLY 278

QY 298 DYAEEMPEIKIKSLAEVAYELVGMKSERVIINWMEIPDWDDPKGRPILLOYARDVRA 357

DB 279 DIMNSLDVKKKLENAVEFG--KTEIADLEAODIKYKMTS-GDBESYIKSRDILH 335

QY 358 TYGLAEKILPFAIQSLVYTGLPIDOVGAMSVPFLEWTLIAAPKKELYPNRVERPEET 417

DB 336 TYFLAEELPMHVELSRMIRIPDDVTRSGRQKQVEMLLSEAHKDELAPNRRM-ADS 394

QY 418 YRGALVLEPLRGVHENIAVLDSSMYPNIMIKYVGPDTLVRPEKCGECGAEAEVGA 477

DB 395 YEGAFVLEPARGLHENYICLDFASMPYSIMISYNISPTLTV--GKDDCNV--AEVGH 450

QY 478 RFRRCPPGFYKTLERLELRKRAEMKKYPPDSPEYRLDEROKALKYANASYGYNG 537

DB 451 KFRKHPDGFPRILKMLIEKRREIKKVMCTLDYNSPEYKLDIKQATLKVLTSFGYTG 510

QY 538 MSGARWYCEGCAKAVTAMGRHLIRTAIINIRKLGKVIYGDTSLPVYDPEKVENFIKI 597

DB 511 WSLARWYCEGCEBATAMGRHPIKTSARIKELGFEVLVYGDTSIFVKRQGLSELEKKE 570

QY 598 IK-----EELGFEIKLEKVRKRLPFTAKKRVAGLLEDRIDIVGEFVARGMCELAK 650

DB 571 VKKLQKLSSEMPQIEIDEYETIFVE-KRVAGLQDORILIVKGLVRRGDMCELAK 629

QY 651 EVOTKVEIVLKTSEVNKAVEYVRKIVKLEBEGKPIKELVIMKTLKSLSEYTTAPPV 710

DB 630 KIQKGVITELIKENKPEKAAEVYGVIEIKKAKIPLSDYITIVKGLTRKPSKTESMQAHV 689

QY 711 VAAKRMLSAGRVSPGDKIGYIVKGGRIISORANPYRMVXD-----PSQIDV 758

DB 690 KAAMKAAKRGVIVYTGSKVGFVTKGVNIGRAPSDLIEDFDEVTIDLDGNKYYIKD 749

QY 759 TYVVDHQIIPALRILVFGITERKLLKASATQKTLPDF 797

DB 750 EYVIDHQVLPVLRILSRFETAEOLKGAEE-QQTLIDAF 787

RESULT 13

AAE13416

ID AAE13416

XX AAE13416: standard; Protein; 774 AA.

XX AAE13416:

XX AAE13416:

XX AAE13416:

XX AAE13416:

XX AAE13416:

XX AAE13416:

XX AAE13416:

XX AAE13416:

XX AAE13416:

XX AAE13416:

XX AAE13416:

XX AAE13416:

XX AAE13416:

XX AAE13416:

XX AAE13416:

XX AAE13416:

XX AAE13416:

XX AAE13416:

XX AAE13416:

XX AAE13416:

XX AAE13416:

XX AAE13416:

Archaeoglobus lithotrophicus DNA polymerase.

DNA polymerase; thermostable; improved fidelity;

Archaeoglobus lithotrophicus.

MO200220735-A2.

14-MAR-2002.

06-SEP-2001; 2001EP-0111032.

11-MAY-2000; 2000JP-0138796.

(TOWN) TOYO BOSEKI KK.

Kuroita T, Kitabayashi M, Ishida Y, Komatsubara H, Nishiya Y,

Kawamura B, Kawamura Y, Imanaka T,

WPI; 2002-012658/02.

Modified heat stable polymerase, useful in polymerase chain reactions,

has specific histidine residue mutated to impart e.g. increased

Fidelity -
Claim 11; Page -; 64pp; English

CC The invention relates to modified thermostable DNA polymerase
CC having 3'-5'-exonuclease activity from *Pyrococcus kodakarensis*.
CC In the modified thermostable DNA polymerase, a specific histidin
CC residue in the wild-type enzyme has been replaced by some other
CC amino acid. The modified thermostable DNA polymerase has enhanced
CC amplification efficiency and/or improved fidelity in PCR reaction
CC and is used as reagent for amplification, extension or mutation
CC template DNA in essentially standard polymerase chain reactions.
CC The present sequence is *Pyrococcus kodakarensis* KOD1 strain
CC mutant DNA polymerase.
CC Note: This sequence is not shown in the specification, but is
CC derived from the thermostable DNA polymerase (AEI3405) shown in
CC page 36-42 of the specification.

Sequence 774 AA;

Query Match	29.8%;	Score 1248.5;	DB 23;	Length 774;
Best Local Similarity	36.8%;	Pred. No. 3.3e-108;		
Matches 305;	Conservative 144;	Mismatches 287;	Indels 93;	Gaps 22;

```

0Y 8 VLDSEYVAVGKPEPOVJIIINGIAENGRVVLIDRSPRPYATALLAGADPKOY----AQIR 63
Db 2 IUDTYITEDGKPVIRIR-KENGSGFKIENDTEPEYATLLKODSALIEVKITIAERHG 60
0Y 64 ALSRKSPLIIGVEDDKRYKGRPREVLIRITVLPFAVEAYEVELVKNVDGVLEADIRF 122
Db 61 TVVTYK-----VEKOKFKLGRPVEVMKLTFTHPQDVPAINDKIREHNAVIDIYEYDIPF 116
0Y 124 AMRYLIDHDLPEFTYRYVEABPLENKMGRVDKYLVKSRPEPLYGEALAPTKLDIRIL 183
Db 117 ARRYLIDKGLV-----PYMGD-----EELKML 138
0Y 184 AFDIEVYSKOSPRPERDPAVIYIAKTDGDEVL-----FLAEGDDKPKIREVEY 235
Db 139 AFDIETLYKEGSEAPL-CPILMITSIADEAGARVITKNVDLPYVDVSTEREMIKRLRIV 197
0Y 236 VKRYDPDIIIVGNNNHFPMPYLLRPARILIGIKLDTREAVGAEPYTSVGH---VSVEGRL 292
Db 198 VAKRODVLITNGNDFPAVLYKKRCCKLGINPALGRD-GSEPKIQRNGDPFAVEVAGRI 256
0Y 293 NVDLIDYAEEMPEIKISLEEAVALGVKKSSERVYINMWEIPIYWDPKPKRPLLQYAR 352
Db 257 HEDLPVIRRTINTLPYTLLEAVYEAVFOGPK-EKYVAE--EITTAMEGENLEBAVAYSM 313
0Y 353 DDVRAVTRGAEKILRPPALQLSYVTCLEPDDOGAMSVGRLRMYLIRAFKKKEIVPNRVE 412
Db 314 EBAKATYELGKKEFLPMEADLSRLIQOSLMDVSRSSSTGMLVEMFLIRKAYERNIEBLAPKPD 373
0Y 413 RPE-----ETYGALIVLEPLRGVHENIIVLPSSMYPNIMIKYNNVGPRTIVRPEBKSGEC 466
Db 374 EKEELARRGOSYGGYVKEBERGUMENIYYLDRSLYSPIIITTHNVSPDLRRE----- 428
0Y 468 GCWE---AEVYKGRERRCPRGPFKTVLRLRLRLRKAENAEKKKPPDPSPEYRLDEROKA 524
Db 427 GCKEYDVAPQVGHRCXDPGPFITPSLDBDLLEBRQKIKKKMKA-TIPDIEKKLDYDORA 486
0Y 525 LKVLNASTGVMGSGARWYCECKAKAVTANGRLIRTAI-NIARKGLKVIYGDTSLF 583
Db 486 IKILANSYGGYGVARAWYCKECSASTVANGREYITWTIKEIEESKGFVIYSDTGPF 545
0Y 584 VTY---DPEKVE---NPKIIEKL--GPEIKLEKYKRLPFBKAKRYVAGLLEDRID 634
Db 546 AITPGADAEVYKKKMAEFKITINACLPALBLEYEGFKRGPFV-TKKKXAVIDEEKIT 604
0Y 635 IVGFAVRGDMCLAEVOTKVEIVLVTSEVNKAVEYVRKIVLELEGKQVIEKUYIMK 694
Db 605 TRGLEIVRDMSEIKETQARVLEALLDGDVYKRAVRIKVEVTEKLSKYEIVPREKVIYHE 666
0Y 695 TISKLEBETTTAHPHVAAKRMLSGHIVSPEDKIGIYIVKGGRISORAMPYMWDPSP 754

```

```

D6 665 QITLBDLKDYGATGPHVAVAKRLARGKIPGTVISYIVKSGSRIQDRAIP -FDEEDPT 7233
OY 755 --QIDTVYVDHDIIPAALRIIVGEGITEKKLKASATGQKLPDPLAKK 801
D6 724 KHKIDAEYIENQVLPVERILRAFGRRKEDLRQKTRQGLSAMLKPK 772

```

RESULT 14

ID AAW77018 standard; protein; 775 AA.

AC AAW77018;

DT 19-NOV-1998 (first entry)

Heat-resistant Pfu DNA synthetase II.

DNA synthetase; enzyme; Pfu, *Pyrococcus furiosus*; functional site;

3 XX

XX

XX

XX
2
11
3
3
3

XX
3
555
1
6
6

XX	30 JAN 1966	0075 0010500
DE		

PR	31-JAN-1997;	97JP-0019248.
PR	31-JAN-1997;	97JP-0019248.

PR 02-DEC-1997; 97JP-0332100.
YY

PA (NISC-) JAPAN SCI & TECHNOLOGY

PI Doi H, Hiraki H, Kanai A;
yy

DR WPI; 1998-437455/37.

PT	Prediction of functional sites
BT	improved protein function

PT	site
VV	

PS Claim 17; pages 46-50; 99pp; Japanese.
xy

CC This represents a heat-resistant Pfu DNA synthetase II enzyme obtained
CC modification of a *Burkholderia* strain. The enzyme was purified by

CC sites using the information obtained by the process of the invention
CC The invention provides a process for providing the following information

CC selected protein in a species for which
CC analysis data is available. The protein

CC the prediction of the functional sites of a protein from genomic or cDNA
analysis data and use of this information to produce methods of immunization

CC functionality by mutation at the functional sites so identified.

Sequence 775 AA;

Query Match	29.8%;	Score 1246.5;	DB 19;	Length 775;
Best Local Similarity	36.5%;	Prod No. 510	108	

Matches 304; Conservative 150; Mismatches 282; Indels 97; Gaps 23

QY 8 VLDSSYEWGKEPVIIMGIAENGERVLI DRSRFPYFYALLAPGADPKQV----AQRI R 63

DB 2 VLDVDYTEEGKPVIRLFK-KENGKFKIEHDRFTFRPYIYALTRDDSKIEEVKKITGERHG 60

64 ALSRPKSPIIGVEDDKRKYFGPRKRVLRIRTVLPEAVREYRELKNDGVEDVLEADIRF 123

DB 61 K1VR---1VDVEKVERKEFGKPRITWKLYLEHPRQDVPTIREKVRHFAVDIFFEYDIPF 116

124 AMRXL DHDLEF

```
Db      117 AKRYLIDGLI-----PMEGE-----EELKIL 138
```

```
Db      117 AKRYLIDGLI-----PMEGE-----EELKIL 138
```

```

QY 184 APDIEVYSKQSPRPDPVIVIAVKTDDGDEVL-----FLAEKDRKREPEVEY 235
DB 139 AFDIEVLYHEGE-EFGKGPITIMISYADENAKVITWKNIDIPYEVAVSSREMIKFLNI 197
QY 236 VKRYDPDITVGVNNHFPWPLYLRARILIGIKLDVTRVGAEPPTSVHGH---VSYPGR 292
DB 198 IREKDPDITVTVNGDSFDPFLAKRAKIGIKLITGRD-GSEPMKQIGMTAVEXKGI 256
QY 293 NVLDVYAEEMPEIKIKSLAEVAYLVGVMKSESVYIMWWEIPYMDPKRLLQYAR 352
DB 257 HDPLVHVTIRITNLPYTLLEAVNEAI-FGKPEKRYAD-BIAAMSGENLERVAKISM 313
QY 353 DVVRATYGLAEKILPFAIQLSVYTGPLDOVGAMSVGRLEMYLIRAFPMKELVNRVE 412
DB 314 EDKATYELGKEFLPMEIQLSRLVGQPLMDVSRSSSTGNLVMPFLIRAYERNENVAENKFS 373
QY 413 RPE-----ETVGAIVLEPLRGVHENIAVLDFSSMYPNIMIKNVGPDTLVPRGKCGE 466
DB 374 EEEYQRLRESYTGFGVPEPEKIMENIVYLDPRALYPSIITITNVGSDPTLANEG----- 428
QY 467 CGCWE-APEVKARFRRCPPGFFKTVLERLLELRKRAEMKCPDPSPEYRLDDEKQAL 525
DB 429 CKRYDIAPOVGHKFCOKIPGIPSLGLHLEEROKIKTKMKE-TODPIEKILLDYRQKAI 487
QY 526 KYLANASYGVGMGSGARWYCRECAKAVTAMGRHLIRTA-INIARKLGLKVIYGDTSLPV 584
DB 488 KILANSFYGYGAKARWYCKEACASVTAMGRKVIELVWKELEKRPFKVLYIDTGLVA 547
QY 585 TYD-----PEKVENPKIKIEELG--PEIKLEKTVRLAFTKAKRYAGLLEDRIDI 635
DB 548 TIRGSESEIKKALEFKIYKINSKLPGLLELEFGTRKGFV-TKQRIAVIDEKQVIT 606
QY 636 VGEVAVRGKCELAKEVOTKVEIVLKTSEVNAKVEYRKIVKLESGKPIELKVIYMT 695
DB 607 RGLLEIVRDMSEIAKETQARVLETLHGQVEEAVIYKVEVIAQKLANVELIPPEKAIYEQ 666
QY 696 LSKRLEEYTTFAPIVNAKRLMAGRVSPBDKIGVIVKGGRIISORAMP---YEMVMD 752
DB 667 ITRPLHEKVAIGPVAHAKKLAAGVAKIKPGVIGIYVIRGDRIRRAIPABEY-----D 722
QY 753 PS--QIDVYVVDHQLIPALRLIGVIGITEKKLAKASATGQKTFDFL-AKKS 802
DB 723 PKGKRYDAVEYIENQVLAVALRILEGYRKEDLRYOKTRQVGLTSLWLNKKS 775

```

RESULT 15

AAE13417 standard; Protein: 774 AA.

AAE13417;

12-FEB-2002 (first entry)

P. kodakarensis KOD1 strain mutant DNA polymerase, H147R.

DNA polymerase; thermostable; 3'-5'-exonuclease; improved fidelity;

amplification; extension; mutation; mutant; mutain.

Pyrococcus kodakarensis.

Synthetic.

Key

FT Msc-difference 147

PN EPI154017-A1.

PD 14-NOV-2001.

PE 08-MAY-2001; 2001EP-0111032.

PR 11-MAY-2000; 2000JP-0138796.

```

XX (TOYM) TOYO BOSEKI KK.
PA Kurota T, Kitanayashi M, Ishida Y, Komatsubara H, Nishiya Y,
PI Kawakami B, Kawamura Y, Imanaka T;
PI WPI; 2002-012658/02.
XX Modified heat stable polymerase, useful in polymerase chain reactions,
XX has specific histidine residue mutated to impart e.g. increased
XX fidelity.
XX Claim 12; Page -1: 64pp; English.
XX The invention relates to modified thermostable DNA polymerase
XX having 3'-5'-exonuclease activity from Pyrococcus kodakarensis.
XX In the modified thermostable DNA polymerase, a specific histidine
XX residue in the wild-type enzyme has been replaced by some other
XX amino acid. The modified thermostable DNA polymerase has enhanced
XX amplification efficiency and/or improved fidelity in PCR reaction
XX and is used as reagent for amplification, extension or mutation of
XX template DNA in essentially standard polymerase chain reactions.
XX The present sequence is Pyrococcus kodakarensis KOD1 strain
XX mutant DNA polymerase.
XX Note: This sequence is not shown in the specification, but is
XX derived from the thermostable DNA polymerase (AAE13405) shown in
XX page 36-42 of the specification.
XX Sequence 774 AA:
XX
XX Query Match 29.7%; Score 1245.5; DB 23; Length 774;
XX Best Local Similarity 36.7%; Pred. No. 6.4e-108;
XX Matches 304; Conservative 145; Mismatches 287; Indels 93; Gaps 22;

```

```

QY 8 VLDSSYVVGKPEQVYIINGIAENGERVVLIDSPFPYFALLAGADPKQV---AQRIR 63
DB 2 ILDTDYITEDGPRVIRIK-KENGEFKIEYDTFEEYVALLKDSAIIEVKKITERRG 60
QY 64 ALSRPSPIIGVEDDKRYFGPRPRVLRITVLPAAVEYBELVQNDVGEVDLEADIRF 123
DB 61 TVVTYVR-----VEKQKFKLGRPVEWKLFTHPQVDAIRDKIRREPAVIDYEYDIPF 116
QY 124 AMRYLLDHDLPFTVYRVEAPEPLENKGPRVDKVLVVSREPLYGALAPTLPLRLT 183
DB 117 AKRYLIDKLV-----PWECD-----BELKVL 138
QY 184 APDIEVYSKQSPRPDPVIVIAVKTDDGDEVL-----FLAEKDRKREPEVEY 235
DB 139 AFDIEVLYHEGEPAE-GPILMISYADESGARVITWKNIDIPYDVAVSTERMIKFLNV 197
QY 236 VKRYDPDITVGVNNHFPWPLYLRARILIGIKLDVTRVGAEPPTSVHGH---VSYPGR 292
DB 198 IREKDPDITVTVNGDSFDPFLAKRAKIGIKLITGRD-GSEPMKQIGMTAVEXKGI 256
QY 293 NVLDVYAEEMPEIKIKSLAEVAYLVGVMKSESVYIMWWEIPYMDPKRLLQYAR 352
DB 257 HDPLVHVTIRITNLPYTLLEAVNEAIFGQPK-EKYVAE-BITTAMETGENLERVARYM 313
QY 353 DVVRATYGLAEKILPFAIQLSVYTGPLDOVGAMSVGRLEMYLIRAFPMKELVNRVE 412
DB 314 EDKATYELGKEFLPMEIQLSRLVGQPLMDVSRSSSTGNLVMPFLIRAYERNENVAENKFS 373
QY 413 RPE-----ETVGAIVLEPLRGVHENIAVLDFSSMYPNIMIKNVGPDTLVPRGKCGC 467
DB 374 EKLARRRQSYEGYVKEPEKIMENIVYLDPRALYPSIITITNVGSDPTLANEG----- 426
QY 468 GCWE---APEVKARFRRCPPGFFKTVLERLLELRKRAEMKCPDPSPEYRLDDEKQAL 524
DB 429 CKRYDIAPOVGHKFCOKIPGIPSLGLHLEEROKIKTKMKE-TIPIRKILLDYRQKAI 485
QY 526 KYLANASYGVGMGSGARWYCRECAKAVTAMGRHLIRTA-INIARKLGLKVIYGDTSLPV 583
DB 488 KILANSFYGYGAKARWYCKEACASVTAMGREYITWTIKEIKGFFVIVSDTDGFF 545

```

Qy 584 VTY---DPEKVE---NFIKIIKEEL--GPEIKLEKXVYKRLFTEAKKRYVAGLLEDGRID 634
Db 546 ATIRGADAEITYKKGAMEFLKTIINAKLPGALELEYEGFYKGFV-TKKYIVIDEBGKIT 604
Qy 635 IVGEAVRGDMCELAKVEQTVVEIVLKTSEVNKAVEYVRKIYKLEBEGKPIEKLYWK 694
Db 605 TRGLEIVRRDMSIYAKETQARVLEALLKDGVEKAVRIYKEVTEKLSKYEVPEKLYIHE 664
Qy 695 TLSKRLBEYTEAPHVVAAKRMLSNAGYRVSFGDKIGYIVYKGGRIISORANPYFMVKDPS 754
Db 665 QITRDLKDYKATGPHVAVAKRLAARGVIRDPGVISYIVLKGSGRIQDRAIP-FDEFDPT 723
Qy 755 --QIDVTYVVDHOLIIPALRIIGYFGITEKKLKASATGCKTLFDFLANK 801
Db 724 KHKYDAEYIENQVLPFAVERILRAFGYRKEDLRYQKTRQVGLSAMLPK 772

arch completed: April 23, 2003, 07:30:23
p time : 46 secs

•
•
•

•

•

•

•

GenCore version 5.1.4.p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 07:30:27 ; Search time 252 Seconds
(without alignments)
255,334 Million cell updates/sec

Title: US-10-034-849-2

Perfect score: 4188
Sequence: 1 MTEVFVTVLSSYEYVWKEP.....LKASATGQKTLFPLAKSK 803

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Number of hits satisfying chosen parameters: 301932

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

Published Applications AA:*

- 1: /cgn2_6/prodata/2/pubppa/US08_NEW_PUB pep:*
- 2: /cgn2_6/prodata/2/pubppa/PCR_NEW_PUB pep:*
- 3: /cgn2_6/prodata/2/pubppa/US06_NEW_PUB pep:*
- 4: /cgn2_6/prodata/2/pubppa/US06_PUBCOMB pep:*
- 5: /cgn2_6/prodata/2/pubppa/US07_NEW_PUB pep:*
- 6: /cgn2_6/prodata/2/pubppa/US07_PUBCOMB pep:*
- 7: /cgn2_6/prodata/2/pubppa/PCR_PUBCOMB pep:*
- 8: /cgn2_6/prodata/2/pubppa/US08_PUBCOMB pep:*
- 9: /cgn2_6/prodata/2/pubppa/US09_NEW_PUB pep:*
- 10: /cgn2_6/prodata/2/pubppa/US09_PUBCOMB pep:*
- 11: /cgn2_6/prodata/2/pubppa/US10_NEW_PUB pep:*
- 12: /cgn2_6/prodata/2/pubppa/US10_PUBCOMB pep:*
- 13: /cgn2_6/prodata/2/pubppa/US60_NEW_PUB pep:*
- 14: /cgn2_6/prodata/2/pubppa/US60_PUBCOMB pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4188	100.0	803	9 US-10-029-382-2	Sequence 2, Appli
2	4188	100.0	803	10 US-09-391-340-4	Sequence 4, Appli
3	4188	100.0	803	10 US-09-348-369-4	Sequence 4, Appli
4	4188	100.0	803	10 US-09-348-369-14	Sequence 14, Appli
5	4188	100.0	803	12 US-10-034-849-2	Sequence 2, Appli
6	4188	100.0	803	12 US-10-034-621-2	Sequence 2, Appli
7	4164	99.4	803	10 US-09-348-369-16	Sequence 16, Appli
8	1563	37.3	788	10 US-09-391-340-6	Sequence 6, Appli
9	1563	37.3	788	10 US-09-348-369-6	Sequence 6, Appli
10	1242.5	29.7	774	10 US-09-852-922-2	Sequence 2, Appli
11	1194	28.5	774	10 US-09-803-165-34	Sequence 34, Appli
12	564.5	13.5	844	9 US-10-027-806-30	Sequence 30, Appli
13	564.5	13.5	844	9 US-10-034-623-30	Sequence 30, Appli
14	564.5	13.5	844	9 US-10-027-801-30	Sequence 30, Appli
15	552	13.2	845	9 US-10-027-806-62	Sequence 62, Appli
16	552	13.2	845	9 US-10-034-623-62	Sequence 62, Appli
17	552	13.2	845	9 US-10-027-801-62	Sequence 62, Appli
18	542.5	13.0	877	10 US-09-391-340-8	Sequence 8, Appli
19	542.5	13.0	877	10 US-09-348-369-8	Sequence 8, Appli

20	506.5	12.1	762	10 US-09-391-340-10	Sequence 10, Appli
21	506.5	12.1	762	10 US-09-348-369-10	Sequence 10, Appli
22	452	10.8	1235	10 US-09-904-065-17	Sequence 17, Appli
23	452	10.8	1240	10 US-09-904-065-4	Sequence 4, Appli
24	452	10.8	1240	10 US-09-904-065-15	Sequence 15, Appli
25	450	10.7	1235	10 US-09-904-065-8	Sequence 8, Appli
26	450	10.7	1235	10 US-09-904-065-10	Sequence 10, Appli
27	450	10.7	1235	10 US-09-904-065-18	Sequence 18, Appli
28	450	10.7	1235	10 US-09-904-065-19	Sequence 19, Appli
29	449	10.7	1235	10 US-09-904-065-6	Sequence 6, Appli
30	449	10.7	1235	10 US-09-904-065-16	Sequence 16, Appli
31	445	10.6	1238	10 US-09-904-065-2	Sequence 2, Appli
32	445	10.6	1238	10 US-09-904-065-14	Sequence 14, Appli
33	368	8.8	1242	10 US-09-904-065-12	Sequence 12, Appli
34	368	8.8	1242	10 US-09-904-065-13	Sequence 13, Appli
35	167	4.0	327	10 US-09-925-300-1209	Sequence 1209, App
36	113.5	2.7	517	12 US-10-062-254-328	Sequence 328, App
37	112.5	2.7	159	9 US-10-055-364-50	Sequence 50, App
38	112.5	2.7	405	9 US-10-102-806-445	Sequence 445, App
39	111.5	2.7	599	10 US-09-815-242-11326	Sequence 11326, A
40	109	2.6	833	9 US-09-940-925A-85	Sequence 85, Appli
41	109	2.6	1211	10 US-09-815-242-11608	Sequence 11608, A
42	106.5	2.5	550	10 US-09-815-242-10499	Sequence 10499, A
43	106	2.5	893	9 US-09-891-332A-9	Sequence 9, Appli
44	105.5	2.5	693	10 US-09-815-242-13233	Sequence 13233, A
45	104.5	2.5	777	10 US-09-801-574-88	Sequence 88, Appli

ALIGNMENTS

RESULT 1
US-10-029-382-2
Sequence 2, Application US/10029382
Patent No. US20020164618A1
GENERAL INFORMATION:
APPLICANT: DIVERSA CORPORATION
APPLICANT: CALLEN, Walter
APPLICANT: MATHER, Eric
TITLE OF INVENTION: ENZYMES HAVING HIGH TEMPERATURE POLYMERASE ACTIVITY AND METHODS
FILE REFERENCE: DIVER1350-2
CURRENT APPLICATION NUMBER: US/10/029,382
CURRENT FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: US/09/656,309
PRIOR FILING DATE: 2000-09-06
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 803
TYPE: PRT
ORGANISM: Pyrobolus fumarius
US-10-029-382-2

Query Match 100.0%; Score 4188; DB 9; Length 803;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 803; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MTEVFVTVLSSYEYVWKEPQVIIMGIAENGERVVLIDRSFRPYFYALAPGADPKOVAQ	60
Db	1	MTEVFVTVLSSYEYVWKEPQVIIMGIAENGERVVLIDRSFRPYFYALAPGADPKOVAQ	60
Qy	61	RIRALSPKSPITIGVEDDKRRYFGRPRVIRITVLEAVREYRELKXNDVGVLEAD	120
Db	61	RIRALSPKSPITIGVEDDKRRYFGRPRVIRITVLEAVREYRELKXNDVGVLEAD	120
Qy	121	IRFMRVLIHDHLPFTWYVEAPLEKMGFVDKYLIVKSPPEPLVYGAALPTKLPDL	180
Db	121	IRFMRVLIHDHLPFTWYVEAPLEKMGFVDKYLIVKSPPEPLVYGAALPTKLPDL	180
Qy	181	RILAFDIEVSKQSPRPERRDPVIVIAVKTDDGDEVLFIEGKODRKPIEFPEYVRYD	240
Db	181	RILAFDIEVSKQSPRPERRDPVIVIAVKTDDGDEVLFIEGKODRKPIEFPEYVRYD	240

Wed Apr 23 08:09:12 2003

us-10-034-849-2.rapb

Page 2

```
OY 241 PDIIIVGNNNHPDMPYLLRRARILGILKLDVTRVGAEPPTTSVHGHSVPGRLNVDLYDA 300
DB 241 PDIIIVGNNNHPDMPYLLRRARILGILKLDVTRVGAEPPTTSVHGHSVPGRLNVDLYDA 300
OY 301 EEMPEIKISLEEVAAEYLGWKKSERVVIINMMEIPDYVDDPKKRPILLQYARDVAVATYG 360
DB 301 EEMPEIKISLEEVAAEYLGWKKSERVVIINMMEIPDYVDDPKKRPILLQYARDVAVATYG 360
OY 361 LAEKILPFAIQLSYVTGGLPDQVGAMSVGFRLIEWLIRAAFKMKELVPRNVERPEETRYG 420
DB 361 LAEKILPFAIQLSYVTGGLPDQVGAMSVGFRLIEWLIRAAFKMKELVPRNVERPEETRYG 420
OY 421 AIVLEPLRGVHENIAVLDPSSMYPNIMIKYVGPDTLVRPGEKCGEGCWEAPEVGRFR 480
DB 421 AIVLEPLRGVHENIAVLDPSSMYPNIMIKYVGPDTLVRPGEKCGEGCWEAPEVGRFR 480
OY 481 RCPGPFKTVLERLLELRKRVRAEMKKYPPDSPERYLLDEROKALKVLNAAISYGMWSG 540
DB 481 RCPGPFKTVLERLLELRKRVRAEMKKYPPDSPERYLLDEROKALKVLNAAISYGMWSG 540
OY 541 ARMYCRECAKAVTAMGSHLIRTAINTARGLGLKYVGTDSLPFTYDPEKVENFIKIIE 600
DB 541 ARMYCRECAKAVTAMGSHLIRTAINTARGLGLKYVGTDSLPFTYDPEKVENFIKIIE 600
OY 601 ELGFEIKLEKYKRLFFTEAKKRYAGLLEDGRIDI VGFPAVRGDMCELAKEVOTKVEIV 660
DB 601 ELGFEIKLEKYKRLFFTEAKKRYAGLLEDGRIDI VGFPAVRGDMCELAKEVOTKVEIV 660
OY 661 LKTSVNRKAVEYVRKIVKELEEGKPIEKLVIWKTLSRLEBYTTEAPHVAAKRMLSSAG 720
DB 661 LKTSVNRKAVEYVRKIVKELEEGKPIEKLVIWKTLSRLEBYTTEAPHVAAKRMLSSAG 720
OY 721 YRVSFGDKIGVIVYKGGGRISQAMPYFMVXPDSQIDVTYYVDHQIIPALRILIGYFGIT 780
DB 721 YRVSFGDKIGVIVYKGGGRISQAMPYFMVXPDSQIDVTYYVDHQIIPALRILIGYFGIT 780
OY 781 EKKLKASATGQKTLFDFLAKSK 803
DB 781 EKKLKASATGQKTLFDFLAKSK 803

RESULT 2
US-09-391-340-4
; Sequence 4, Application US/09391340A
; Patent No. US20020013455A1
; GENERAL INFORMATION:
; APPLICANT: Callen, Walter
; APPLICANT: Mather, Eric
; TITLE OF INVENTION: ISOLATION AND IDENTIFICATION OF NOVEL POLYMERASES
; FILE REFERENCE: 09010/027001
; CURRENT APPLICATION NUMBER: US/09/391,340A
; EARLIER FILING DATE: 1999-09-07
; EARLIER APPLICATION NUMBER: US 08/907,166
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 803
; TYPE: PRT
; ORGANISM: Pyrolobus fumarius
US-09-391-340-4

Query Match 100.0%, Score 4188; DB 10; Length 803;
Best Local Similarity 100.0%, Pred. No. 0;
Matches 803; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MTEVFTVLDSSYEYVKGPEOYIINGIAENGERVLLIDRSFRPYFALLAPADPKQVAG 60
DB 1 MTEVFTVLDSSYEYVKGPEOYIINGIAENGERVLLIDRSFRPYFALLAPADPKQVAG 60
OY 61 RIRALSRPKSPITIGVEDDKRYFGRRPRVLIRITVLEAVRERELKRVNDGVEVLEAD 120
DB 61 RIRALSRPKSPITIGVEDDKRYFGRRPRVLIRITVLEAVRERELKRVNDGVEVLEAD 120
```

```
DB 61 RIRALSRPKSPITIGVEDDKRYFGRRPRVLIRITVLEAVRERELKRVNDGVEVLEAD 120
OY 121 IRPAMRYLIDHDLPFTMYVEAEAPLENKMGGRVUKYLVKSRREPIYXGALA.PTGLPDL 180
DB 121 IRPAMRYLIDHDLPFTMYVEAEAPLENKMGGRVUKYLVKSRREPIYXGALA.PTGLPDL 180
OY 181 RIIAPFIEYYSKQSPREPERDPIVIAVKTDDGEVLFIAEKODRPIEFYEVYKRYD 240
DB 181 RIIAPFIEYYSKQSPREPERDPIVIAVKTDDGEVLFIAEKODRPIEFYEVYKRYD 240
OY 241 PDIIIVGNNNHPDMPYLLRRARILGILKLDVTRVGAEPPTTSVHGHSVPGRLNVDLYDA 300
DB 241 PDIIIVGNNNHPDMPYLLRRARILGILKLDVTRVGAEPPTTSVHGHSVPGRLNVDLYDA 300
OY 301 EEMPEIKISLEEVAAEYLGWKKSERVVIINMMEIPDYVDDPKKRPILLQYARDVAVATYG 360
DB 301 EEMPEIKISLEEVAAEYLGWKKSERVVIINMMEIPDYVDDPKKRPILLQYARDVAVATYG 360
OY 361 LAEKILPFAIQLSYVTGGLPDQVGAMSVGFRLIEWLIRAAFKMKELVPRNVERPEETRYG 420
DB 361 LAEKILPFAIQLSYVTGGLPDQVGAMSVGFRLIEWLIRAAFKMKELVPRNVERPEETRYG 420
OY 421 AIVLEPLRGVHENIAVLDPSSMYPNIMIKYVGPDTLVRPGEKCGEGCWEAPEVGRFR 480
DB 421 AIVLEPLRGVHENIAVLDPSSMYPNIMIKYVGPDTLVRPGEKCGEGCWEAPEVGRFR 480
OY 481 RCPGPFKTVLERLLELRKRVRAEMKKYPPDSPERYLLDEROKALKVLNAAISYGMWSG 540
DB 481 RCPGPFKTVLERLLELRKRVRAEMKKYPPDSPERYLLDEROKALKVLNAAISYGMWSG 540
OY 541 ARMYCRECAKAVTAMGSHLIRTAINTARGLGLKYVGTDSLPFTYDPEKVENFIKIIE 600
DB 541 ARMYCRECAKAVTAMGSHLIRTAINTARGLGLKYVGTDSLPFTYDPEKVENFIKIIE 600
OY 601 ELGFEIKLEKYKRLFFTEAKKRYAGLLEDGRIDI VGFPAVRGDMCELAKEVOTKVEIV 660
DB 601 ELGFEIKLEKYKRLFFTEAKKRYAGLLEDGRIDI VGFPAVRGDMCELAKEVOTKVEIV 660
OY 661 LKTSVNRKAVEYVRKIVKELEEGKPIEKLVIWKTLSRLEBYTTEAPHVAAKRMLSSAG 720
DB 661 LKTSVNRKAVEYVRKIVKELEEGKPIEKLVIWKTLSRLEBYTTEAPHVAAKRMLSSAG 720
OY 721 YRVSFGDKIGVIVYKGGGRISQAMPYFMVXPDSQIDVTYYVDHQIIPALRILIGYFGIT 780
DB 721 YRVSFGDKIGVIVYKGGGRISQAMPYFMVXPDSQIDVTYYVDHQIIPALRILIGYFGIT 780
OY 781 EKKLKASATGQKTLFDFLAKSK 803
DB 781 EKKLKASATGQKTLFDFLAKSK 803

RESULT 3
US-09-948-369-4
; Sequence 4, Application US/09948369
; Patent No. US20020132243A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: CALLEN, Walter
; APPLICANT: MATHER, Eric
; TITLE OF INVENTION: ENZYMES HAVING HIGH TEMPERATURE POLYMERASE ACTIVITY AND METHODS
; FILE REFERENCE: DIVER1350-3
; CURRENT APPLICATION NUMBER: US/09/948,369
; EARLIER FILING DATE: 2001-09-06
; EARLIER APPLICATION NUMBER: US 09/556,309
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: US 09/391,340
; PRIOR FILING DATE: 1999-09-07
; PRIOR APPLICATION NUMBER: US 08/907,166
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
```


LENGTH: 803
 TYPE: PRT
 ORGANISM: Pyrobolus fumarius
 US-09-948-369-4

Query Match 100.0%; Score 4188; DB 10; Length 803;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 803; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTEVFTVLDSSEYVNGKEPOVITWGIAENGERVYLIDRSFRPFYALLAGADPKOVAQ 60
 DB 1 MTEVFTVLDSSEYVNGKEPOVITWGIAENGERVYLIDRSFRPFYALLAGADPKOVAQ 60
 QY 61 RIRALSRPKSPIIGVEDDKRYFGRPRVRIRITVLEAAREYRELKAVNDGVEDVLEAD 120
 DB 61 RIRALSRPKSPIIGVEDDKRYFGRPRVRIRITVLEAAREYRELKAVNDGVEDVLEAD 120
 QY 121 IRFAMRYLIDHDLFPFTWYRVEAEPLNKGFRVDKYLVSRRPEPLYGALAPTKLPDL 180
 DB 121 IRFAMRYLIDHDLFPFTWYRVEAEPLNKGFRVDKYLVSRRPEPLYGALAPTKLPDL 180
 QY 181 RILAFDIEVYSKQSPREBDPVIVIAVKTDDGDEVLFIAEGDKDRKPIREFVEYVKRYD 240
 DB 181 RILAFDIEVYSKQSPREBDPVIVIAVKTDDGDEVLFIAEGDKDRKPIREFVEYVKRYD 240
 QY 241 PDIIIVGNHNFDMFYLRRARILIGIKLDTVRVGAEPITSVGHVSVGRANVDLYDYA 300
 DB 241 PDIIIVGNHNFDMFYLRRARILIGIKLDTVRVGAEPITSVGHVSVGRANVDLYDYA 300
 QY 301 EEMBEIKIKSLEEAVALYGVKKSERVIINWMEIPDYMDPKRPLLOVARDVRAATYG 360
 DB 301 EEMBEIKIKSLEEAVALYGVKKSERVIINWMEIPDYMDPKRPLLOVARDVRAATYG 360
 QY 361 LAEKILPFAILOSTVGTPLDQVAMSGFRLEWYLIRAAFKMKELVNRVERPEETRYG 420
 DB 361 LAEKILPFAILOSTVGTPLDQVAMSGFRLEWYLIRAAFKMKELVNRVERPEETRYG 420
 QY 421 AIVLEPLRGVHENIAVLDSSMYPNIMIKYVGPDTLVRPGEKGCCEGCEWAPVKRFR 480
 DB 421 AIVLEPLRGVHENIAVLDSSMYPNIMIKYVGPDTLVRPGEKGCCEGCEWAPVKRFR 480
 QY 481 RCPGPFKTVLERLLELRKVRAMKCYPPDSPEYRLDEROKALKVLANASYGMWSG 540
 DB 481 RCPGPFKTVLERLLELRKVRAMKCYPPDSPEYRLDEROKALKVLANASYGMWSG 540
 QY 541 ARWYCRECAKAVTAMGRHLIRTAIINARKLGLKVIYGDTSLFTTYDEKVENFIKIIE 600
 DB 541 ARWYCRECAKAVTAMGRHLIRTAIINARKLGLKVIYGDTSLFTTYDEKVENFIKIIE 600
 QY 601 ELGFEIKLEKYKRLFTFEAKKRYAGLLEDGRIDIVGFEARGDMCELAKEVQTKVEIV 660
 DB 601 ELGFEIKLEKYKRLFTFEAKKRYAGLLEDGRIDIVGFEARGDMCELAKEVQTKVEIV 660
 QY 661 LKTSSEVNAVEYVARKIYKELEEGKVPLEKVIWKTLSKREYTTAEAPHVAAKRLMSAG 720
 DB 661 LKTSSEVNAVEYVARKIYKELEEGKVPLEKVIWKTLSKREYTTAEAPHVAAKRLMSAG 720
 QY 721 YRVSFGDKIGIVYVGGGRISQRAWPFYMWKDSQIDVTYVVDHQIIPALRLILGYGKIT 780
 DB 721 YRVSFGDKIGIVYVGGGRISQRAWPFYMWKDSQIDVTYVVDHQIIPALRLILGYGKIT 780
 QY 781 EKKLKASATGQKTLFDFLAKSK 803
 DB 781 EKKLKASATGQKTLFDFLAKSK 803

RESULT 4
 US-09-948-369-14
 Sequence 14; Application US/09948369
 Patent No. US20020132243A1
 GENERAL INFORMATION:
 APPLICANT: DIVERSA CORPORATION
 APPLICANT: CALLEN, Walter

APPLICANT: MATHUR, Eric
 TITLE OF INVENTION: ENZYMES HAVING HIGH TEMPERATURE POLYMERASE ACTIVITY AND METHODS
 TITLE OF INVENTION: THEREOF
 FILE REFERENCE: DIVER1350-3
 CURRENT APPLICATION NUMBER: US/09/948,369
 PRIOR FILING DATE: 2001-09-06
 PRIOR APPLICATION NUMBER: US 09/656,309
 PRIOR FILING DATE: 2000-09-06
 PRIOR APPLICATION NUMBER: US 09/391,340
 PRIOR FILING DATE: 1999-09-07
 PRIOR APPLICATION NUMBER: US 08/907,166
 PRIOR FILING DATE: 1997-08-06
 NUMBER OF SEQ ID NOS: 16
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 14
 LENGTH: 803
 TYPE: PRT
 ORGANISM: Pyrobolus fumarius
 US-09-948-369-14

Query Match 100.0%; Score 4188; DB 10; Length 803;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 803; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTEVFTVLDSSEYVNGKEPOVITWGIAENGERVYLIDRSFRPFYALLAGADPKOVAQ 60
 DB 1 MTEVFTVLDSSEYVNGKEPOVITWGIAENGERVYLIDRSFRPFYALLAGADPKOVAQ 60
 QY 61 RIRALSRPKSPIIGVEDDKRYFGRPRVRIRITVLEAAREYRELKAVNDGVEDVLEAD 120
 DB 61 RIRALSRPKSPIIGVEDDKRYFGRPRVRIRITVLEAAREYRELKAVNDGVEDVLEAD 120
 QY 121 IRFAMRYLIDHDLFPFTWYRVEAEPLNKGFRVDKYLVSRRPEPLYGALAPTKLPDL 180
 DB 121 IRFAMRYLIDHDLFPFTWYRVEAEPLNKGFRVDKYLVSRRPEPLYGALAPTKLPDL 180
 QY 181 RILAFDIEVYSKQSPREBDPVIVIAVKTDDGDEVLFIAEGDKDRKPIREFVEYVKRYD 240
 DB 181 RILAFDIEVYSKQSPREBDPVIVIAVKTDDGDEVLFIAEGDKDRKPIREFVEYVKRYD 240
 QY 241 PDIIIVGNHNFDMFYLRRARILIGIKLDTVRVGAEPITSVGHVSVGRANVDLYDYA 300
 DB 241 PDIIIVGNHNFDMFYLRRARILIGIKLDTVRVGAEPITSVGHVSVGRANVDLYDYA 300
 QY 301 EEMBEIKIKSLEEAVALYGVKKSERVIINWMEIPDYMDPKRPLLOVARDVRAATYG 360
 DB 301 EEMBEIKIKSLEEAVALYGVKKSERVIINWMEIPDYMDPKRPLLOVARDVRAATYG 360
 QY 361 LAEKILPFAILOSTVGTPLDQVAMSGFRLEWYLIRAAFKMKELVNRVERPEETRYG 420
 DB 361 LAEKILPFAILOSTVGTPLDQVAMSGFRLEWYLIRAAFKMKELVNRVERPEETRYG 420
 QY 421 AIVLEPLRGVHENIAVLDSSMYPNIMIKYVGPDTLVRPGEKGCCEGCEWAPVKRFR 480
 DB 421 AIVLEPLRGVHENIAVLDSSMYPNIMIKYVGPDTLVRPGEKGCCEGCEWAPVKRFR 480
 QY 481 RCPGPFKTVLERLLELRKVRAMKCYPPDSPEYRLDEROKALKVLANASYGMWSG 540
 DB 481 RCPGPFKTVLERLLELRKVRAMKCYPPDSPEYRLDEROKALKVLANASYGMWSG 540
 QY 541 ARWYCRECAKAVTAMGRHLIRTAIINARKLGLKVIYGDTSLFTTYDEKVENFIKIIE 600
 DB 541 ARWYCRECAKAVTAMGRHLIRTAIINARKLGLKVIYGDTSLFTTYDEKVENFIKIIE 600
 QY 601 ELGFEIKLEKYKRLFTFEAKKRYAGLLEDGRIDIVGFEARGDMCELAKEVQTKVEIV 660
 DB 601 ELGFEIKLEKYKRLFTFEAKKRYAGLLEDGRIDIVGFEARGDMCELAKEVQTKVEIV 660
 QY 661 LKTSSEVNAVEYVARKIYKELEEGKVPLEKVIWKTLSKREYTTAEAPHVAAKRLMSAG 720
 DB 661 LKTSSEVNAVEYVARKIYKELEEGKVPLEKVIWKTLSKREYTTAEAPHVAAKRLMSAG 720
 QY 721 YRVSFGDKIGIVYVGGGRISQRAWPFYMWKDSQIDVTYVVDHQIIPALRLILGYGKIT 780

Wed Apr 23 08:09:12 2003

us-10-034-849-2.rapb

Page 4

Db 721 YRVSPGDKIGYIVYVGGGRISORAMPYFVWKDPSQIDVTYVDHQQIIPALRILGYFGIT 780
Qy 781 EKKLKASATGOKTLFDFLAKSK 803
Db 781 EKKLKASATGOKTLFDFLAKSK 803

RESULT 5

US-10-034-849-2
Sequence 2, Application US/10034849
Patent No. US2002015108A1
GENERAL INFORMATION:
APPLICANT: DIVERSA CORPORATION
APPLICANT: CALLEN, Walter
APPLICANT: MATHER, Eric
TITLE OF INVENTION: ENZYMES HAVING HIGH TEMPERATURE POLYMERASE ACTIVITY
TITLE OF INVENTION: AND METHODS OF USE
FILE REFERENCE: DIVER1350-2
CURRENT APPLICATION NUMBER: US/10/034, 849
CURRENT FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: US/09/656,309
PRIOR FILING DATE: 2000-09-06
NUMBER OF SEQ ID NOS: 2
SOFTWARE: Patentin version 3.0
SEQ ID NO 2
LENGTH: 803
TYPE: PRF
ORGANISM: Pyrollobus fumarius
US-10-034-849-2

Query Match 100.0%; Score 4188; DB 12; Length 803;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 803; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTEVFTVLDSSYEYVGEKPOVITNGIAENGERVVLIDRSFRPFYFALLAPGADPKQVQ 60
Db 1 MTEVFTVLDSSYEYVGEKPOVITNGIAENGERVVLIDRSFRPFYFALLAPGADPKQVQ 60
Qy 61 RIRALSRPKSPIIGVEDDKRYTGRPRVLRIRTVLPEAVREYRELKVNVDGEVDLEAD 120
Db 61 RIRALSRPKSPIIGVEDDKRYTGRPRVLRIRTVLPEAVREYRELKVNVDGEVDLEAD 120
Qy 121 IRFAMRYLIDHDLFPFTWYRVEAEPLNKGFRVDKYLKSRPEPLYGELAPTKLPDL 180
Db 121 IRFAMRYLIDHDLFPFTWYRVEAEPLNKGFRVDKYLKSRPEPLYGELAPTKLPDL 180
Qy 181 RILAFDIEVYKSGSPRPDPVIVIAVKTDDGDEVLFIAEGKDDKPIREVEYVKRYD 240
Db 181 RILAFDIEVYKSGSPRPDPVIVIAVKTDDGDEVLFIAEGKDDKPIREVEYVKRYD 240
Qy 241 PDIIYGVNNHFMWPLRABRILGKIDVTRVGAEPITSVGHVSGRLAVDLXYA 300
Db 241 PDIIYGVNNHFMWPLRABRILGKIDVTRVGAEPITSVGHVSGRLAVDLXYA 300
Qy 301 EEMPEIKIKSLAEVAAYLGVKKSERVITNNWEIPDYWDPKKRPILLQYARDVRYATYG 360
Db 301 EEMPEIKIKSLAEVAAYLGVKKSERVITNNWEIPDYWDPKKRPILLQYARDVRYATYG 360
Qy 361 LAEKILPFAIQLSYVTGLPLDVGAMSVGFLEWYLLRAAFKMKELVPRNVERPREETYG 420
Db 361 LAEKILPFAIQLSYVTGLPLDVGAMSVGFLEWYLLRAAFKMKELVPRNVERPREETYG 420
Qy 421 AIVLEPLRGVHENIAVLDFSSMYPNIMIKYVGPDLVVRGEGCGCGMEAPVYKHF 480
Db 421 AIVLEPLRGVHENIAVLDFSSMYPNIMIKYVGPDLVVRGEGCGCGMEAPVYKHF 480
Qy 481 RCPGPFKTVLERLLELRGRVRAEMKKYPPDSPEYRLDLDRQALVLANBYGIMGSG 540
Db 481 RCPGPFKTVLERLLELRGRVRAEMKKYPPDSPEYRLDLDRQALVLANBYGIMGSG 540
Qy 541 ARWYCRCAKAVTAMGRHILRTAINIAKGLKXIVYDQDLSLVYTDPEKVENFIKIKE 600

Db 541 ARWYCRCAKAVTAMGRHILRTAINIAKGLKXIVYDQDLSLVYTDPEKVENFIKIKE 600
Qy 601 ELGFEIKLEKRYKRLFTTEAKRYAGLBDGIDIVGEAVRGWCELANEVOTKVEIV 660
Db 601 ELGFEIKLEKRYKRLFTTEAKRYAGLBDGIDIVGEAVRGWCELANEVOTKVEIV 660
Qy 661 LKTSVNAVAYVRYKIVLEEGKVPIDKVIWKTLSKRLSEYTTTEAPHYAAKRLSAG 720
Db 661 LKTSVNAVAYVRYKIVLEEGKVPIDKVIWKTLSKRLSEYTTTEAPHYAAKRLSAG 720
Qy 721 YRVSPGDKIGYIVYVGGGRISORAMPYFVWKDPSQIDVTYVDHQQIIPALRILGYFGIT 780
Db 721 YRVSPGDKIGYIVYVGGGRISORAMPYFVWKDPSQIDVTYVDHQQIIPALRILGYFGIT 780
Qy 781 EKKLKASATGOKTLFDFLAKSK 803
Db 781 EKKLKASATGOKTLFDFLAKSK 803

RESULT 6

US-10-034-621-2
Sequence 2, Application US/10034621
Patent No. US20020132397A1
GENERAL INFORMATION:
APPLICANT: DIVERSA CORPORATION
APPLICANT: CALLEN, Walter
APPLICANT: MATHER, Eric
TITLE OF INVENTION: ENZYMES HAVING HIGH TEMPERATURE POLYMERASE ACTIVITY
TITLE OF INVENTION: AND METHODS OF USE
FILE REFERENCE: DIVER1350-2
CURRENT APPLICATION NUMBER: US/10/034, 621
CURRENT FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: US/09/656,309
PRIOR FILING DATE: 2000-09-06
NUMBER OF SEQ ID NOS: 2
SOFTWARE: Patentin version 3.0
SEQ ID NO 2
LENGTH: 803
TYPE: PRF
ORGANISM: Pyrollobus fumarius
US-10-034-621-2

Query Match 100.0%; Score 4188; DB 12; Length 803;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 803; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTEVFTVLDSSYEYVGEKPOVITNGIAENGERVVLIDRSFRPFYFALLAPGADPKQVQ 60
Db 1 MTEVFTVLDSSYEYVGEKPOVITNGIAENGERVVLIDRSFRPFYFALLAPGADPKQVQ 60
Qy 61 RIRALSRPKSPIIGVEDDKRYTGRPRVLRIRTVLPEAVREYRELKVNVDGEVDLEAD 120
Db 61 RIRALSRPKSPIIGVEDDKRYTGRPRVLRIRTVLPEAVREYRELKVNVDGEVDLEAD 120
Qy 121 IRFAMRYLIDHDLFPFTWYRVEAEPLNKGFRVDKYLKSRPEPLYGELAPTKLPDL 180
Db 121 IRFAMRYLIDHDLFPFTWYRVEAEPLNKGFRVDKYLKSRPEPLYGELAPTKLPDL 180
Qy 181 RILAFDIEVYKSGSPRPDPVIVIAVKTDDGDEVLFIAEGKDDKPIREVEYVKRYD 240
Db 181 RILAFDIEVYKSGSPRPDPVIVIAVKTDDGDEVLFIAEGKDDKPIREVEYVKRYD 240
Qy 241 PDIIYGVNNHFMWPLRABRILGKIDVTRVGAEPITSVGHVSGRLAVDLXYA 300
Db 241 PDIIYGVNNHFMWPLRABRILGKIDVTRVGAEPITSVGHVSGRLAVDLXYA 300
Qy 301 EEMPEIKIKSLAEVAAYLGVKKSERVITNNWEIPDYWDPKKRPILLQYARDVRYATYG 360
Db 301 EEMPEIKIKSLAEVAAYLGVKKSERVITNNWEIPDYWDPKKRPILLQYARDVRYATYG 360
Qy 361 LAEKILPFAIQLSYVTGLPLDVGAMSVGFLEWYLLRAAFKMKELVPRNVERPREETYG 420

```

Db 361 LAEKILPFAIQLSVYTGPLPDQVGAMSVGRLEWYLIRAAFKMELVNNRVERPEETRG 420
Qy 421 AIVLEPLRGHENAVLDFSSMTYPMIKYVGPDLTVRPGKCGCCGCEWAPVKHFR 480
Db 421 AIVLEPLRGHENAVLDFSSMTYPMIKYVGPDLTVRPGKCGCCGCEWAPVKHFR 480
Qy 481 RCPGPFKTVLERLELKRRAEMKKYPPDSPERYLLDEROKALVLANASYGMGMSG 540
Db 481 RCPGPFKTVLERLELKRRAEMKKYPPDSPERYLLDEROKALVLANASYGMGMSG 540
Qy 541 ARMYCRECAKAVTAMGRHLIRTAIINARKLGLKVIYGDTSLEFVYDPEKXENFIKIIE 600
Db 541 ARMYCRECAKAVTAMGRHLIRTAIINARKLGLKVIYGDTSLEFVYDPEKXENFIKIIE 600
Qy 601 ELGFEIKLEKYYKRLFPEAKKRYAGLLEDGRIDIVGEFVARGDWCCLAKEVQTKVEIV 660
Db 601 ELGFEIKLEKYYKRLFPEAKKRYAGLLEDGRIDIVGEFVARGDWCCLAKEVQTKVEIV 660
Qy 661 LKTSBVNKAVEYVYRKIVKELEGGKVPLEKLVIMKTLSKRLEEYTEAPHVVAARKMSAG 720
Db 661 LKTSBVNKAVEYVYRKIVKELEGGKVPLEKLVIMKTLSKRLEEYTEAPHVVAARKMSAG 720
Qy 721 YRVSPGDKIGYVIYKGGGRISQRAPIYFVMDPSQIDVTYYVDHQIIPALRIILGYFGIT 780
Db 721 YRVSPGDKIGYVIYKGGGRISQRAPIYFVMDPSQIDVTYYVDHQIIPALRIILGYFGIT 780
Qy 781 EKKLKASATGQKTLFDFLAKSK 803
Db 781 EKKLKASATGQKTLFDFLAKSK 803

```

RESULT 7

```

US-09-948-369-16
; Sequence 16, Application US/09948369
; Patent No. US20020132243A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: CALLEN, Walter
; APPLICANT: MATHEN, Eric
; TITLE OF INVENTION: ENZYMES HAVING HIGH TEMPERATURE POLYMERASE ACTIVITY AND METHODS C
; FILE REFERENCE: DIVERL350-3
; CURRENT APPLICATION NUMBER: US/09/948,369
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: US 09/656,309
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: US 09/391,340
; PRIOR FILING DATE: 1999-09-07
; PRIOR APPLICATION NUMBER: US 08/907,166
; PRIOR FILING DATE: 1997-08-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
; LENGTH: 803
; TYPE: PRT
; ORGANISM: Pyrococcus furiosus
US-09-948-369-16

```

```

Query Match 99.4%; Score 4164; DB 10; Length 803;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 797; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 1 MTEVFTVLDSSEYVVGKEPQVIIMGIAENGERVVLIDRSFRPYFYALLAPGADPKQVAQ 60
Db 1 MTEVFTVLDSSEYVVGKEPQVIIMGIAENGERVVLIDRSFRPYFYALLAPGADPKQVAQ 60
Qy 61 RIRALSRKSPITIGVEDDKKRYFGRRVLRIRTLPEAVREYRLVKNVNGVEVLEAD 120
Db 61 RIRALSRKSPITIGVEDDKKRYFGRRVLRIRTLPEAVREYRLVKNVNGVEVLEAD 120
Qy 121 IRFAARVYLIDHDLFPFTYVRAEABLENMGRVVDKVLVSRPPLVGEALAPTKLPDL 180
Db 121 IRFAARVYLIDHDLFPFTYVRAEABLENMGRVVDKVLVSRPPLVGEALAPTKLPDL 180

```

```

Db 121 IRFAARVYLIDHDLFPFTYVRAEABLENMGRVVDKVLVKSREPLVGEALAPTKLPDL 180
Qy 181 RILAFDIEVYSKQSGPPEEDPVIIVIAVKTDDGDEVLFIAEGKODKPIREPVEYVKRYD 240
Db 181 RILAFDIEVYSKQSGPPEEDPVIIVIAVKTDDGDEVLFIAEGKODKPIREPVEYVKRYD 240
Qy 241 PDIIYGNHNFHMPYLLRBRARILIGIKLDVTRRGAPPTSVGHVSVPRLANVDLYDA 300
Db 241 PDIIYGNHNFHMPYLLRBRARILIGIKLDVTRRGAPPTSVGHVSVPRLANVDLYDA 300
Qy 301 EEMPEIKISLEAEVAYELGVYKKSERYIIMWEIPDWDPPKRPPLLOQYARDDVRATYG 360
Db 301 EEMPEIKISLEAEVAYELGVYKKSERYIIMWEIPDWDPPKRPPLLOQYARDDVRATYG 360
Qy 361 LAEKILPFAIQLSVYTGPLPDQVGAMSVGRLEWYLIRAAFKMELVNNRVERPEETRG 420
Db 361 LAEKILPFAIQLSVYTGPLPDQVGAMSVGRLEWYLIRAAFKMELVNNRVERPEETRG 420
Qy 421 AIVLEPLRGHENAVLDFSSMTYPMIKYVGPDLTVRPGKCGCCGCEWAPVKHFR 480
Db 421 AIVLEPLRGHENAVLDFSSMTYPMIKYVGPDLTVRPGKCGCCGCEWAPVKHFR 480
Qy 481 RCPGPFKTVLERLELKRRAEMKKYPPDSPERYLLDEROKALVLANASYGMGMSG 540
Db 481 RCPGPFKTVLERLELKRRAEMKKYPPDSPERYLLDEROKALVLANASYGMGMSG 540
Qy 541 ARMYCRECAKAVTAMGRHLIRTAIINARKLGLKVIYGDTSLEFVYDPEKXENFIKIIE 600
Db 541 ARMYCRECAKAVTAMGRHLIRTAIINARKLGLKVIYGDTSLEFVYDPEKXENFIKIIE 600
Qy 601 ELGFEIKLEKYYKRLFPEAKKRYAGLLEDGRIDIVGEFVARGDWCCLAKEVQTKVEIV 660
Db 601 ELGFEIKLEKYYKRLFPEAKKRYAGLLEDGRIDIVGEFVARGDWCCLAKEVQTKVEIV 660
Qy 661 LKTSBVNKAVEYVYRKIVKELEGGKVPLEKLVIMKTLSKRLEEYTEAPHVVAARKMSAG 720
Db 661 LKTSBVNKAVEYVYRKIVKELEGGKVPLEKLVIMKTLSKRLEEYTEAPHVVAARKMSAG 720
Qy 721 YRVSPGDKIGYVIYKGGGRISQRAPIYFVMDPSQIDVTYYVDHQIIPALRIILGYFGIT 780
Db 721 YRVSPGDKIGYVIYKGGGRISQRAPIYFVMDPSQIDVTYYVDHQIIPALRIILGYFGIT 780
Qy 781 EKKLKASATGQKTLFDFLAKSK 803
Db 781 EKKLKASATGQKTLFDFLAKSK 803

```

RESULT 8

```

US-09-391-340-6
; Sequence 6, Application US/09391340A
; Patent No. US20020013455A1
; GENERAL INFORMATION:
; APPLICANT: Calten, Walter
; APPLICANT: Mathen, Eric
; TITLE OF INVENTION: ISOLATION AND IDENTIFICATION OF NOVEL POLYMERASES
; FILE REFERENCE: 09010/027001
; CURRENT APPLICATION NUMBER: US/09/391,340A
; PRIOR FILING DATE: 1999-09-07
; PRIOR APPLICATION NUMBER: US 08/907,166
; PRIOR FILING DATE: 1997-08-06
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 788
; TYPE: PRT
; ORGANISM: Archaeoglobus fulgidus
US-09-391-340-6

```

```

Query Match 37.3%; Score 1563; DB 10; Length 788;
Best Local Similarity 41.3%; Pred. No. 5; 7e-120;
Matches 338; Conservative 154; Mismatches 273; Indels 54; Gaps 18;

```

```

Qy 1 MTEVFTVLDSSEYVVGKEPQVIIMGIAENGERVVLIDRSFRPYFYALLAPGADPKQVAQ 60

```

Wed Apr 23 08:09:12 2003

us-10-034-849-2.rapb

Page 6

```
Db 1 MIKVGWLLDADYITENDRAVIRLWCKDEEG-IFIAVDSFQYPYALKEEGITAEIDYK 59
Qy 61 -RIRALSRPKSPIIGVEDDKRYFGRPRVLRIRTVLPEAVREYRELVKAVDVEA 119
Db 60 IKVQTKKEVITP-LKVEETAKNLGREVVEFKIYARHQVHPKLEEVSO---YLEIREA 115
Qy 120 DIRFARMYLIHDHLPFTWYRVEAEPLNKGFRVDKYLVK--SRPELYGALAPYKL 177
Db 116 DIPFAYRYLIDKNLACMDGVILEGVERREK-GLRCYEIKRIERDSRD-----F 163
Qy 178 PDRLIAPDIEVYKSGSPRPDPVYIAVKTDDGDEVFLAEGKODRKPIREFVEYK 237
Db 164 PELKWAAPDCMLSEVGMDEPDKDPIIVISISGEYEIL---NGDNRELTRFVKIIR 220
Qy 238 RYDPDIIVGYNNHFPWYLLRARILOIKLDVTRVGAETTSVGHVSPGRLVNDLY 297
Db 221 DIDPDIIVGYNDSFDPWPIIKRAEKLKVKLDIGR--DRSELAI RGRPKIAGRLVNDLY 278
Qy 298 DYAEEMPEIKISLEAEVYLGVMKKSERYIINWMEIPDYWDDPKRPLLQYARDVRA 357
Db 279 DIAMRSIDVKKYKLENAVEFLG--KKEIADIEBAKDIYKHWTS--GDRESVIKYSRODILH 335
Qy 358 TYGLAEKILPFAIOLSVYTGPLDPOVAMSVGRLEWYLIIRAAFKMKELVPRVERPEET 417
Db 336 TYPIAEELLPMHYELSRMIRIPLDDVTRSGRGKQVEWLLSEAHKLGELAPNPREM-ADS 394
Qy 418 YRGAIVLEPLRGVHENIAVLDFSSMYPNIMIKYVGPDTLVRPGEKGCSCGCEAPVEYK 477
Db 395 YEGAFVLEBPARGHENVICLDFASMYPSIMISPDITVI--GKDDCNV--APEVGH 450
Qy 478 RFRCPGPFKTVLERLELRKVRAMKXKPPDPSPRYLLDEROKALVYLANASGYNG 537
Db 451 KFRGHPDGFRRILKOLIEKREIKCKMKTLDNPSPEYKLDIKQATLKVLTNSFYGTG 510
Qy 538 WSGARWYCECAKAVTANGRHILRTAINIARKLGKVIYGDTSLVTVYDPEKVENFIKI 597
Db 511 WSLARNYCKEACATTANGRHIFITSARIKELGEVLYGDTDSIFVKKDGSLSELKE 570
Qy 598 IK-----BELGFEIKLEKRYKRLFTFAKKRYAGLLEDGRIDIVGEFAVGDCELA 650
Db 571 VKKLIGLSEEMPIQIEIDYEYETIFVE-KKRYAGLTDGRITIVGLVGRGDCELA 629
Qy 651 EVQTKVEIVLKTSEVNKAVEYRKIVLEEGVPIEKLVIMKTLSEYTTAPHV 710
Db 630 KIQGVIEIILKEKNPEKAAYKGVIEIKAKIPLDEYIYKGLTRKPSKYESMAHV 689
Qy 711 VAAKRLASAGYRSPGDKIGYIVKGGRIISORAMPYFMYKD-----PSQIDV 758
Db 690 KAAMKAARAGIYVITGSKGVFVTKGVNIGBRAFPDLIEDPDGEVITDLDGNKYKIDK 749
Qy 759 TYVVDHIIIPALRIIGYFGITEKKUKASATGOKTLFDF 797
Db 750 EYVIDHQVLPSTVRLIRLBRFGYTEADLKGAAB-QOTLDAF 787
```

RESULT 9
US-09-948-369-6
Sequence 6, Application US/09948369
Patent No. US20020132243A1
GENERAL INFORMATION:
APPLICANT: DIVERSA CORPORATION
APPLICANT: CALLEN, Walter
APPLICANT: MAYHUR, Eric
TITLE OF INVENTION: ENZYMES HAVING HIGH TEMPERATURE POLYMERASE ACTIVITY AND METHODS C
FILE REFERENCE: DIVER1350-3
CURRENT APPLICATION NUMBER: US/09/948,369
CURRENT FILING DATE: 2001-09-06
PRIOR APPLICATION NUMBER: US 09/656,309
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: US 09/391,340
PRIOR FILING DATE: 1999-09-07

PRIOR APPLICATION NUMBER: US 08/907,166
PRIOR FILING DATE: 1997-08-06
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.0
SEQ ID NO 6
LENGTH: 788
ORGANISM: Archaeoglobus fulgidus
US-09-948-369-6

Query Match 37.3%; Score 1563; DB 10; Length 788;
Best Local Similarity 41.3%; Fred. No. 5.7e-120;
Matches 338; Conservative 154; Mismatches 273; Indels 54; Gaps 18;

```
Qy 1 MTEVFTVLDSSYEYVGEKPOVYIINGIAENGERVVLIDRSFRPYFALLAPGADPKVQA 60
Db 1 MIKVGWLLDADYITENDRAVIRLWCKDEEG-IFIAVDSFQYPYALKEEGITAEIDYK 59
Qy 61 -RIRALSRPKSPIIGVEDDKRYFGRPRVLRIRTVLPEAVREYRELVKAVDVEA 119
Db 60 IKVQTKKEVITP-LKVEETAKNLGREVVEFKIYARHQVHPKLEEVSO---YLEIREA 115
Qy 120 DIRFARMYLIHDHLPFTWYRVEAEPLNKGFRVDKYLVK--SRPELYGALAPYKL 177
Db 116 DIPFAYRYLIDKNLACMDGVILEGVERREK-GLRCYEIKRIERDSRD-----F 163
Qy 178 PDRLIAPDIEVYKSGSPRPDPVYIAVKTDDGDEVFLAEGKODRKPIREFVEYK 237
Db 164 PELKWAAPDCMLSEVGMDEPDKDPIIVISISGEYEIL---NGDNRELTRFVKIIR 220
Qy 238 RYDPDIIVGYNNHFPWYLLRARILOIKLDVTRVGAETTSVGHVSPGRLVNDLY 297
Db 221 DIDPDIIVGYNDSFDPWPIIKRAEKLKVKLDIGR--DRSELAI RGRPKIAGRLVNDLY 278
Qy 298 DYAEEMPEIKISLEAEVYLGVMKKSERYIINWMEIPDYWDDPKRPLLQYARDVRA 357
Db 279 DIAMRSIDVKKYKLENAVEFLG--KKEIADIEBAKDIYKHWTS--GDRESVIKYSRODILH 335
Qy 358 TYGLAEKILPFAIOLSVYTGPLDPOVAMSVGRLEWYLIIRAAFKMKELVPRVERPEET 417
Db 336 TYPIAEELLPMHYELSRMIRIPLDDVTRSGRGKQVEWLLSEAHKLGELAPNPREM-ADS 394
Qy 418 YRGAIVLEPLRGVHENIAVLDFSSMYPNIMIKYVGPDTLVRPGEKGCSCGCEAPVEYK 477
Db 395 YEGAFVLEBPARGHENVICLDFASMYPSIMISPDITVI--GKDDCNV--APEVGH 450
Qy 478 RFRCPGPFKTVLERLELRKVRAMKXKPPDPSPRYLLDEROKALVYLANASGYNG 537
Db 451 KFRGHPDGFRRILKOLIEKREIKCKMKTLDNPSPEYKLDIKQATLKVLTNSFYGTG 510
Qy 538 WSGARWYCECAKAVTANGRHILRTAINIARKLGKVIYGDTSLVTVYDPEKVENFIKI 597
Db 511 WSLARNYCKEACATTANGRHIFITSARIKELGEVLYGDTDSIFVKKDGSLSELKE 570
Qy 598 IK-----BELGFEIKLEKRYKRLFTFAKKRYAGLLEDGRIDIVGEFAVGDCELA 650
Db 571 VKKLIGLSEEMPIQIEIDYEYETIFVE-KKRYAGLTDGRITIVGLVGRGDCELA 629
Qy 651 EVQTKVEIVLKTSEVNKAVEYRKIVLEEGVPIEKLVIMKTLSEYTTAPHV 710
Db 630 KIQGVIEIILKEKNPEKAAYKGVIEIKAKIPLDEYIYKGLTRKPSKYESMAHV 689
Qy 711 VAAKRLASAGYRSPGDKIGYIVKGGRIISORAMPYFMYKD-----PSQIDV 758
Db 690 KAAMKAARAGIYVITGSKGVFVTKGVNIGBRAFPDLIEDPDGEVITDLDGNKYKIDK 749
Qy 759 TYVVDHIIIPALRIIGYFGITEKKUKASATGOKTLFDF 797
Db 750 EYVIDHQVLPSTVRLIRLBRFGYTEADLKGAAB-QOTLDAF 787
```

RESULT 10
US-09-852-922-2

Db 488 AVKHAASYGVMGCRKAWMSKECAESVTAMGRHYIEMTIKEIKERGFVLYADTGC 547
 QY 583 FVYDPRKVENFIKIIKEELG-----EIKLEKYKRLPFTAKKRYVGLLEDRI 633
 Db 548 YATIPGEPETIKKAKPEFKTINSKLPGLLELEBEPYJNGFFV-AKKRYAVIDEEGR 606
 QY 634 DIVPEAVRGDMCELAKEVQTRVVEIVLKTSEVNAVEYRKIVKELEEGKVPLEKVIW 693
 Db 607 TTRGLEVVRDMSKIAKEVQAVLALIKEDSVKAVEIVKDVVEIACYQVPEKLVYH 666
 QY 694 KTLSEKLEEYTTAPPHVAAKMLSAGRYVSPGDKIGYIVKGGRIISORAMPYVNDP 753
 Db 667 EOIITKDLSEYKAIQPHVAIAKRLAKGIVKPGTIIISYIVLRSGSKISDRV-ILTSEYDP 725
 QY 754 S-QIDVTYVVDHIIIPALRLILGYFGITKEKKLAKSATGOKTLPDEIAR 800
 Db 726 KKHKYPDYYIENQVLPVILRILEAFGRKEDLKYGSSKQGVDAWLKK 774

RESULT 12

Sequence 30, Application US/10027806
 Patent No. US20020160476A1
 GENERAL INFORMATION:

APPLICANT: Swanson, Ronald V.
 APPLICANT: Feldman, Robert A.
 APPLICANT: Schleper, Christa
 TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
 FILE REFERENCE: DCOIP.002A
 CURRENT APPLICATION NUMBER: US/10/027,806
 CURRENT FILING DATE: 2001-12-21
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/408,020
 PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-29
 NUMBER OF SEQ ID NOS: 123
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 30
 LENGTH: 844
 TYPE: PRT
 ORGANISM: Cenarchaeum symbiosum
 US-10-027-806-30

Query Match 13.5%; Score 564.5; DB 9; Length 844;
 Best Local Similarity 25.9%; Pred. No. 8.3e-38;
 Matches 192; Conservative 127; Mismatches 300; Indels 122; Gaps 25;

Db 108 KNVGVEVDLEADIRFAMRYLIDHDLFPFTWYRVA-----EPLNKMGFRVDRKVVLY 160
 Db 106 KSIINIMDTWESDIKYENYLVDKSLVGRYYSVSGKVIPLHMPISDEV-----KLMK 160
 Db 161 KSRPELYGEALAPTK-----LPDLRIAPDIENVYSKQG--SPRPERD 201
 Db 161 SLIMDKVVDGMDRKEFRREFIAGWADLNPRIPIRLSFDIEVDEEGRIIPDKISDR 220
 QY 202 PYVIAVKTDDGDEVLPFT-----AEKDKRKPIREFVEYVRYRDP-----DIIV 245
 Db 221 RVTAVGPAATDGLKQVFLRSGAESEBENVITGVEVYVYDEADMIRDALSVIGSYPFL 280
 QY 246 GYNNHFMWYLLRARRILGIR-LDVTIRVGAEPITSVHG-HVSV--GRLANVDLYDVA 300
 Db 281 TYNDGDFDPMYMLNRRARRLGVSDDIPLYYMRDSATLRHGVHLDLRTFSNRSFOLYFA 340
 QY 301 EEMPEIKIKSLEVAEYLVGMKSERVIINMWEIPDYWDPKRPPILL-----QYARDVR 356
 Db 341 AKYDYSLNSVTKA--MLGEGK-----VDY--GVKGLDITLYQTANYCYHDAR 384
 QY 357 ATYGLA-----EKLPPAIOLSVYVGLPLDOVGANSVGRLEWYLIAPAKMEIYPNVE 412
 Db 385 LTLELSTFGNEILMDLLVTSRIAMPIDMSRMGVSQWIRSLYYEHRQNALIPRDE 444
 QY 413 R-----PEETRGAIYVEPLRGVHENIAVLDSSSMYPRIMAKKNVSGDTLYR 459
 Db 445 LEGSRREVSNDAYIKDKKFRGLVVEBEGIHFDVYVMDPASILIKRYNLSLEYV-- 502

QY 460 PGEKC--GEGCWEAPEYKHFRRCPFPFKTVLRLLELRKVRABMKYPPDSPEYRL 517
 Db 503 ---RCHVACECKKNTIPDNIHWCTKNNGLTSMIISLSDLEVNYYKSLSTSITBE--- 556
 QY 518 LDERQ-----KALKVLNANSYGMGSGARWYCEKCAVATANGRLIIRTAIINARKUG 571
 Db 557 --ORQOYTVISQALKVNLNANSYGMGAELFPPLYPAAEATTAAGRYIINQITISHEQWG 614
 QY 572 LKVIYDSDLSLVYTYDE--KVENFIKIIKEELGEIKLEKYKRLPFTAKKRYAGLLE 629
 Db 615 VRVLGDDTSLPIK-DEBERQIHEIVEHAKKEHGELEVDDEKERYVVLNRRKNVFGYTR 673
 QY 630 DGRIDIVGEAVRGDMCELAKEVQTRVVEI--VLTSEVNAKAVEYRKIV--KELE 682
 Db 674 AGKVDKGLTGKSHTPPFIKELFYSLDILSGVESEDEPESAKRISKALACGRLE 733
 QY 683 GKVPLEKLVIMKTLSEKLEEYTTAPPHVAAKMLSAGRYVSPGDKIGYIVKGGRIISQ 742
 Db 734 RQIPLVLDLAFNMISKAPSEYVTVPOHIRARILLENAREVYKGDIIISYKVNANKGVK- 792
 QY 743 RAMPYRMKDSQIDVYTYVD 763
 Db 793 ---PVENAR-AGEVDTSKYLE 809

RESULT 13

US-10-034-623-30
 Sequence 30, Application US/10034623
 Publication No. US20020198365A1
 GENERAL INFORMATION:

APPLICANT: Swanson, Ronald V.
 APPLICANT: Feldman, Robert A.
 APPLICANT: Schleper, Christa
 TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
 FILE REFERENCE: DCOIP.002A
 CURRENT APPLICATION NUMBER: US/10/034,623
 CURRENT FILING DATE: 2001-12-21
 PRIOR APPLICATION NUMBER: 09/408,020
 PRIOR FILING DATE: 1999-09-29
 PRIOR APPLICATION NUMBER: 60/102,294
 PRIOR FILING DATE: 1998-09-29
 NUMBER OF SEQ ID NOS: 123
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 30
 LENGTH: 844
 TYPE: PRT
 ORGANISM: Cenarchaeum symbiosum
 US-10-034-623-30

Query Match 13.5%; Score 564.5; DB 9; Length 844;
 Best Local Similarity 25.9%; Pred. No. 8.3e-38;
 Matches 192; Conservative 127; Mismatches 300; Indels 122; Gaps 25;

QY 108 KNVGVEVDLEADIRFAMRYLIDHDLFPFTWYRVA-----EPLNKMGFRVDRKVVLY 160
 Db 106 KSIINIMDTWESDIKYENYLVDKSLVGRYYSVSGKVIPLHMPISDEV-----KLMK 160
 Db 161 KSRPELYGEALAPTK-----LPDLRIAPDIENVYSKQG--SPRPERD 201
 Db 161 SLIMDKVVDGMDRKEFRREFIAGWADLNPRIPIRLSFDIEVDEEGRIIPDKISDR 220
 QY 202 PYVIAVKTDDGDEVLPFT-----AEKDKRKPIREFVEYVRYRDP-----DIIV 245
 Db 221 RVTAVGPAATDGLKQVFLRSGAESEBENVITGVEVYVYDEADMIRDALSVIGSYPFL 280
 QY 246 GYNNHFMWYLLRARRILGIR-LDVTIRVGAEPITSVHG-HVSV--GRLANVDLYDVA 300
 Db 281 TYNDGDFDPMYMLNRRARRLGVSDDIPLYYMRDSATLRHGVHLDLRTFSNRSFOLYFA 340
 QY 301 EEMPEIKIKSLEVAEYLVGMKSERVIINMWEIPDYWDPKRPPILL-----QYARDVR 356
 Db 341 AKYDYSLNSVTKA--MLGEGK-----VDY--GVKGLDITLYQTANYCYHDAR 384

QY 357 ATYGLA-----EKILPFAIOLSVYTGCLPLDOVGAMSVGFRLBWLIRAAFMKELVPRVE 412
 Db 385 LTLELSTFGNEILMDLIVVTSRIARMPIDDMSRMGVQWISLSLYEHRONALIPRDE 444
 QY 413 R-----PEETYGAIIVLEPLRGVHENIAVLDPSSMYPNIMIKNVGPDTLVR 459
 Db 445 LEGRSREVSNDVAIKDKKFRGGLVVEPEEGIHFDVTAMDFAISLYPSIIKRNLSYETV-- 502
 QY 460 PGEKC--GEGGCWAPBEVKRFRCPGPFKTVLERLELKRRAEMKYPDPSPERYL 517
 Db 503 ---RCVHAECCKNTIPDTHNVCTKNNGLTSMITGSLDLRNVYKSLKSTSTEE-- 556
 QY 518 LDERO-----KALKVLANASYGYMGSGARWYCRECAKAVTAMGRHLIRTAIINARKLG 571
 Db 557 --OQOQVTVISQALKVLANASYGVWGAIFPLFLPAEATTAAGRYYIIMOTISHCEOMG 614
 QY 572 LKVIYGDTSLFVTVYDPE--KVENFIKIKBELFEIKLEKVKRYLFTTEAKRYAGLLE 629
 Db 615 VRVLYGDTDSLFIK--DPERQIHEIVEHAKKEHVELEVDKEXRYVVLNKRKNFVTR 673
 QY 630 DGRIDIVGFEAVRGDWCLAKEVOTKVEI---VLKTSVNKAVEYRKIV---KELEE 682
 Db 674 AGKDVVKGLTGKKSHTPPFIKELFYSLLDILSGVSEDEFESAKRISKALAACGRLEE 733
 QY 683 GKVDIEKLVIMKTLSKRLEETTEAPHVNAAKRMLAGRYVSPGDKIGYIVYKGGRIQ 742
 Db 734 RQILVLDLAFVNMISKAPSEYKTVPOHIRARLLENAREYKGDIIISYVNMKTGVK- 792
 QY 743 RAMPYFVWKDPSQIDVTYYVD 763
 Db 793 ---EVEMAR--AGEVDTSKYLE 809

RESULT 14
 US-10-027-801-30
 ; Sequence 30, Application US/10027801
 ; Publication No. US20030054364A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Swanson, Ronald V.
 ; APPLICANT: Feldman, Robert A.
 ; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
 ; FILE REFERENCE: DCOIP.002A
 ; CURRENT APPLICATION NUMBER: US/10/027,801
 ; PRIOR FILING DATE: 2001-12-21
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/408,020
 ; NUMBER OF SEQ ID NOS: 123
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 30
 ; LENGTH: 844
 ; TYPE: PRF
 ; ORGANISM: Cenarchaeum symbiosum
 US-10-027-801-30

Query Match 13.5%; Score 564.5; DB 9; Length 844;
 Best Local Similarity 25.9%; Pred. No. 8.3e-38;
 Matches 192; Conservative 127; Mismatches 300; Indels 122; Gaps 25;
 QY 108 KNVGDVEDVLEADIRFAMRYLIDHDLFPFTWYRVEA-----EPLNMGFRVVDKVLV 160
 Db 106 KSRINIMTWESDIKYENLYDKSLVGRYYSVSGKVIPIHDMISDEV-----KALK 160
 QY 161 KSRPEPLGEGALAPTK-----LPLRLIAPDIEVYSKOG--SPRPERD 201
 Db 161 SLMDKVVDEGADKKEFRFETAGWADLLNCPTRIRRLSFEIIVDSSEGRIPDKISDR 220
 QY 202 FYIVAVATDGDDEVLFI-----AEGKDRKPIREFVYVYKYP-----DIIV 245
 Db 221 RVTAVGPAATDGLKQVFLVRSAGEGEGVTVGVEVFDKADWIRDALSVTSGYPFVL 280
 QY 246 GYNNNHFMFPYLLRRARILGIK-LDVTFRVGAEPPTSVMG--HVSVP---GRILVLDLYYA 300

Db 281 TYNGDDFDMYMLNRRARLGVSDSDIPLVMMRDSATLRHGHLIDYRTFSNRSPQLYFA 340
 QY 301 EEMPEIKIKSIEVAEVLGVWKKSERVIIMWELPDYWDPKKRPILL---QYARDVR 356
 Db 341 AKYTDYSINSVYKA--MLGEGK-----VDY--GVKGLDLYTOTANVCYHAR 384
 QY 357 ATYGLA-----EKILPFAIOLSVYTGCLPLDOVGAMSVGFRLBWLIRAAFMKELVPRVE 412
 Db 385 LTLELSTFGNEILMDLIVVTSRIARMPIDDMSRMGVQWISLSLYEHRONALIPRDE 444
 QY 413 R-----PEETYGAIIVLEPLRGVHENIAVLDPSSMYPNIMIKNVGPDTLVR 459
 Db 445 LEGRSREVSNDVAIKDKKFRGGLVVEPEEGIHFDVTAMDFAISLYPSIIKRNLSYETV-- 502
 QY 460 PGEKC--GEGGCWAPBEVKRFRCPGPFKTVLERLELKRRAEMKYPDPSPERYL 517
 Db 503 ---RCVHAECCKNTIPDTHNVCTKNNGLTSMITGSLDLRNVYKSLKSTSTEE-- 556
 QY 518 LDERO-----KALKVLANASYGYMGSGARWYCRECAKAVTAMGRHLIRTAIINARKLG 571
 Db 557 --OQOQVTVISQALKVLANASYGVWGAIFPLFLPAEATTAAGRYYIIMOTISHCEOMG 614
 QY 572 LKVIYGDTSLFVTVYDPE--KVENFIKIKBELFEIKLEKVKRYLFTTEAKRYAGLLE 629
 Db 615 VRVLYGDTDSLFIK--DPERQIHEIVEHAKKEHVELEVDKEXRYVVLNKRKNFVTR 673
 QY 630 DGRIDIVGFEAVRGDWCLAKEVOTKVEI---VLKTSVNKAVEYRKIV---KELEE 682
 Db 674 AGKDVVKGLTGKKSHTPPFIKELFYSLLDILSGVSEDEFESAKRISKALAACGRLEE 733
 QY 683 GKVDIEKLVIMKTLSKRLEETTEAPHVNAAKRMLAGRYVSPGDKIGYIVYKGGRIQ 742
 Db 734 RQILVLDLAFVNMISKAPSEYKTVPOHIRARLLENAREYKGDIIISYVNMKTGVK- 792
 QY 743 RAMPYFVWKDPSQIDVTYYVD 763
 Db 793 ---EVEMAR--AGEVDTSKYLE 809

RESULT 15
 US-10-027-806-62
 ; Sequence 62, Application US/10027806
 ; Patent No. US20020160476A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Swanson, Ronald V.
 ; APPLICANT: Feldman, Robert A.
 ; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
 ; FILE REFERENCE: DCOIP.002A
 ; CURRENT APPLICATION NUMBER: US/10/027,806
 ; PRIOR FILING DATE: 2001-12-21
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/408,020
 ; NUMBER OF SEQ ID NOS: 123
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 62
 ; LENGTH: 845
 ; TYPE: PRF
 ; ORGANISM: Cenarchaeum symbiosum
 US-10-027-806-62

Query Match 13.2%; Score 552; DB 9; Length 845;
 Best Local Similarity 26.0%; Pred. No. 8.9e-37;
 Matches 209; Conservative 133; Mismatches 339; Indels 124; Gaps 27;
 QY 42 RPYFYALLAPGADPKQVAGRIARLSRPSPIGVEDDKRYKFGRR--VLRIITVLEA 99
 Db 47 KPYCYTRLPSE-----LGFGRVEDVLGIGQVNRHDLIADKVPVSKITVSDPLA 97
 QY 100 VREYRELIVKNVGDVEDVLEADIRFAMRYLIDHDLFPFTWYRV--EAPLENMGFRVVDK 156
 Db 98 ICGTHS-EKSRINVIDWESDIKYENLYDAGLVGRYYSVSGEVIPIHDMISDEV-K 155

Wed Apr 23 08:09:12 2003

us-10-034-849-2.rapb

Page 10

QY 157 VYLVKSRPEPLYGEALAPTK-----LPDDRILAPDIEVYSKOG-----SP 196
DB 156 IALKSLIMDLIDEGWADRKEFRERFIAGMADLNLQIPRIRRLSPDIEVDSSEGRIPDAK 215
QY 197 RPERDPYIVIAVKTDDGDEVLFA-----EGDDBKPIREFEVYKRPDI-----244
DB 216 ISDRRYTAVGFAATDGLRVLVLSKGADEGANBTPBGEVVFYDEKXENDMRDALIIG 275
QY 245 -----VGYNNHEDWYLLRARIIGI-KLDVTRVGAEPITSG-HVSV-GRNAV 294
DB 276 SYPFVLTYNGDDPDMYNNRRARLGVADSDIPLYMRDSATLRGHVLDLRTFSNRF 335
QY 295 DLYDYAEMPEIKIKSLIEVAVYLGWKKSERVIINWELPDYMDPKRPPLLOYARD 354
DB 336 QLYAFAAKYTDYSLNSVSA-MDEGKVDYVSLG--DLTY-----OTANYCYHD 383
QY 355 VRATYGLA-----EKILPFAIOLSYVTGLPLDOVGANSVGFRLMVLIRAAFOMKELVNR 410
DB 384 ARLTLELSTFGNEILMDLIVTSRIARMPIDMSRMGVSQMIRSLIYEHROQNALIPRR 443
DB 411 VER-----PEETYRGAIVLEPLRGVENIAVADPFSMYPNIMIKVNGPDTL 457
DB 444 DELEKRSQVSNDAVVKDKKFRGGLVVEPEGIHPDVTVMDFASLYPSIIKYRNLSEYEV 503
QY 458 --VAPGEKCGECGCWEAPFVKRFRRCPPGFVTLERLLELRKVRBAEMKKYPPDSPEY 515
DB 504 RCVHP-----ECKXNTIPTNNHWCTKXNGILTSWILGSLDLAVNYKSLSKQSITEE 557
QY 516 RLUDERQ-----KALKVLANASYGWMGSGARWYCHRECAKATANGRLITAINARK 569
DB 558 ----QROQYVISQALKVLANASYGWAELIFPLYFPAHEATTAVGRITIMQIISHCEQ 613
QY 570 LGLKVIYGDPSLFTYDPE--XVENFIKIKEELGFEIKLEKYKLEFTEAKRYAGL 627
DB 614 MGVAVLYGDDLSLPIK-NPEERQIHDIYEHAKKEHVELEVDKEYRYVVLNRRKNYFGV 672
QY 628 LEDGRIDYGEFAVRGDMCELAKEVQTVVEI--VLKTSVNRKAVEYRKIV---KEL 680
DB 673 TKSGLVAVKGLTGKSKHTPPFKELFYSGLDILSAVQTEDEFESAKLKISKAIAAGKRL 732
QY 681 EERGVPLEKLVWKLSKRLEETTEAPHVVAARMLSAGYRVSPODKIGYIV--KGGG 738
DB 733 EERGVPLADLAFNWMISKAPSEYKTVPOHIRARLEENAREYKGDIIISYKVMNKTV 792
QY 739 RISQRAWPYFMVDPQIDVTYYVD 763
DB 793 KPVEMA-----QAGEVDTSKYLE 810

Arch completed: April 23, 2003, 07:46:23
db time : 255 secs

ALIGNMENTS

RESULT 1

B56277 DNA-directed DNA polymerase (EC 2.7.7.7) II - Pyrodicticum occultum
 C>Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 20-Jun-2000
 C/Accession: B56277
 R/Unmott, T.; Ishino, Y.; Doi, H.; Kato, I.
 J/Bacteriol. 177, 2164-2177, 1995
 A>Title: The hyperthermophilic archaeon Pyrodicticum occultum has two alpha-like DNA poly
 A/Reference number: A56277; MUID:95238290; PMID:7721707
 A/Accession: B56277
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-803 <DEM>
 A/Cross-references: GB:D36574; NID:9807829; PIDN:BA07580.1; PID:9807830
 C/Superfamily: herpesvirus DNA-directed DNA polymerase
 C/Keywords: nucleotidyltransferase

Query Match 3.5%; Score 287 DB 2; Length 803;
 Best Local Similarity 100.0%; Pred. No. 2.6e-19;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 626 GLEEDGRIDVGFPAVRGDMCLAKEVO 653
 DB 628 GLEEDGRIDVGFPAVRGDMCLAKEVO 655

RESULT 2

JC7382 DNA-directed DNA polymerase (EC 2.7.7.7) B3 - Sulfurisphaera ohwakuensis
 N/Alternate names: DNA polymerase alpha, beta, gamma; DNA polymerase I, II, III
 C/Species: Sulfurisphaera ohwakuensis
 C/Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 31-Dec-2000
 C/Accession: JC7382
 R/Riwal, T.; Kurosawa, N.; Itch, Y.H.; Kimura, N.; Horinuchi, T.
 DNA Res. 7, 243-251, 2000
 A>Title: Sequence analysis of three family B DNA polymerases from the thermoacidophilic
 A/Reference number: JC7380
 A/Accession: JC7382
 A/Molecule type: DNA
 A/Residues: 1-781 <IMA>
 A/Cross-references: DDBJ:AB032376
 C/Comment: This enzyme has both 3'-5' exonuclease and polymerase activities, and plays a
 C/Genetics:
 A/Gene: B3
 C/Keywords: DNA replication; exonuclease; metal binding; nucleotidyltransferase

Query Match 1.7%; Score 14; DB 2; Length 781;
 Best Local Similarity 100.0%; Pred. No. 3.5e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 633 IDIVGFPAVRGDMC 646
 DB 617 IDIVGFPAVRGDMC 630

RESULT 3

E72515 Probable DNA-directed DNA polymerase APE2098 - Aeropyrum pernix (strain KI)
 C/Species: Aeropyrum pernix
 C/Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
 C/Accession: E72515
 R/Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Halkawa, Y.; Jin-no, K.; Takah
 awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
 DNA Res. 6, 83-101, 1999
 A>Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
 A/Reference number: A72450; MUID:99310339; PMID:10382966
 A/Accession: E72515
 A/Status: preliminary
 A/Molecule type: DNA

A/Residues: 1-784 <KAW>
 A/Cross-references: DDBJ:AP000063; NID:95105654; PIDN:BA081109.1; PID:95105797
 A/Experimental source: strain KI
 C/Genetics:
 A/Gene: APE2098
 C/Superfamily: herpesvirus DNA-directed DNA polymerase

Query Match 1.7%; Score 14; DB 2; Length 784;
 Best Local Similarity 100.0%; Pred. No. 3.5e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 449 KYNVGPDTLVPRGE 462
 DB 433 KYNVGPDTLVPRGE 446

RESULT 4

JC5186 DNA-directed DNA polymerase (EC 2.7.7.7) - Sulfolobus acidocaldarius
 C/Species: Sulfolobus acidocaldarius
 C/Date: 20-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 18-Jun-1999
 C/Accession: JC5186; PC4255
 R/Drumstovill, N.; Pokholok, D.; Lottepeich, F.; Prangishvili, D.; Rechin, V.
 Gene 177, 271-273, 1996
 A>Title: The DNA polymerase-encoding gene from a thermoacidophilic archaeon Sulfolobus a
 A/Reference number: JC5186; MUID:97080536; PMID:8921861
 A/Accession: JC5186
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-875 <DAT1>
 A/Cross-references: GB:U33846; NID:91685081; PIDN:AA044598.1; PID:91685082
 A/Accession: PC4255
 A/Status: preliminary
 A/Molecule type: protein
 A/Residues: 713-726/777-790 <DAT2>
 C/Comment: This enzyme belongs to family B of DNA polymerase, and has proof reading acti
 C/Superfamily: herpesvirus DNA-directed DNA polymerase
 C/Keywords: nucleotidyltransferase
 F/223-235/Region: exonuclease signature
 F/286-300/Region: exonuclease signature
 F/402-414/Region: exonuclease signature

Query Match 1.5%; Score 12; DB 2; Length 875;
 Best Local Similarity 100.0%; Pred. No. 0.004;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 572 LKVIYGDPSLF 583
 DB 646 LKVIYGDPSLF 657

RESULT 5

T41940 DNA polymerase - human herpesvirus 7 (strain J1)
 C/Species: human herpesvirus 7
 A/Variety: strain J1
 C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
 C/Accession: T41940
 R/Nicholas, J.
 submitted to the EMBL Data Library, December 1995
 A/Description: Determination and analysis of the complete nucleotide sequence of human
 A/Reference number: 222022
 A/Accession: T41940
 A/Status: preliminary; translated from GB/EMBL/DDBJ
 A/Molecule type: DNA
 A/Residues: 1-1012 <NIC>
 A/Cross-references: EMBL:U03400; PIDN:AA054700.1
 A/Experimental source: strain J1
 C/Genetics:
 A/Note: U38
 C/Superfamily: herpesvirus DNA-directed DNA polymerase

Query Match 1.5%; Score 12; DB 2; Length 1012;

Best Local Similarity 100.0%; Pred. No. 0.0046;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 574 VIYGDTSLEFV 585
|||||

DB 735 VIYGDTSLEFV 746
|||||

RESULT 6

S55603
DNA polymerase replicative chain - equine herpesvirus 2

C/Species: equine herpesvirus 2
C/Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 18-Jun-1999

C/Accession: S55603
R/Albrecht, E.A.R.; Watson, M.S.; Aird, H.C.; Perry, J.; Davison, A.J.

J. Mol. Biol. 249, 520-528, 1995

A/Title: The DNA sequence of equine herpesvirus 2.

A/Reference number: S55594; MUID:95302501; PMID:7783207

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA
A/Residues: 1-1008 <TEL>

A/Cross-references: GB:020824; NID:G695172; PIDN:AAQ13796.1; PID:G695181

A/Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1995

C/Superfamily: herpesvirus DNA-directed DNA polymerase

Query Match
Best Local Similarity 1.4%; Score 11; DB 2; Length 1008;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 573 KVIYGDTSLEFV 583
|||||

DB 744 KVIYGDTSLEFV 754
|||||

RESULT 7

DJBEM2
DNA-directed DNA polymerase (EC 2.7.7.7) - saimiriine herpesvirus 1 (strain 11)

C/Species: saimiriine herpesvirus 1

A/Note: host Saimiri sciureus (common squirrel monkey)

C/Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 11-Jun-1999

C/Accession: I36806
R/Albrecht, J.

submitted to the EMBL Data Library, January 1992

A/Description: Primary structure of the herpesvirus saimiri genome.

A/Reference number: A36806

A/Accession: I36806

A/Molecule type: DNA
A/Residues: 1-1009 <ALB>

A/Cross-references: GB:X64346; NID:G60320; PIDN:CAA45632.1; PID:G60330

R/Albrecht, J.C.; Nicholas, J.; Biller, D.; Cameron, K.R.; Biesinger, B.; Newman, C.; W.

J. Virol. 66, 5047-5058, 1992

A/Title: Primary structure of the herpesvirus saimiri genome.

A/Reference number: A37309; MUID:9233688; PMID:131287

A/Contents: annotation; possible protein-coding frames

A/Note: neither amino acid nor nucleotide sequence is given

C/Genetics:

A/Gene: 9

C/Superfamily: herpesvirus DNA-directed DNA polymerase

C/Keywords: DNA binding; DNA replication; nucleotidyltransferase

Query Match
Best Local Similarity 1.4%; Score 11; DB 1; Length 1009;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 574 VIYGDTSLEFV 584
|||||

DB 737 VIYGDTSLEFV 747
|||||

RESULT 8

T42925
DNA-directed DNA polymerase (EC 2.7.7.7) - ateline herpesvirus 3 (strain 73)

C/Species: ateline herpesvirus 3

A/Variety: strain 73

C/Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 18-Feb-2000

C/Accession: T42925

R/Albrecht, J.C.; Fleckenstein, B.

submitted to the EMBL Data Library, August 1998

A/Description: Primary structure of the herpesvirus atelines genome.

A/Reference number: Z22274

A/Accession: T42925

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA
A/Residues: 1-1009 <ALB>

A/Cross-references: EMBL:AF083424; PIDN:AAQ95533.1

A/Experimental source: strain 73

C/Superfamily: herpesvirus DNA-directed DNA polymerase

C/Keywords: nucleotidyltransferase

Query Match
Best Local Similarity 1.4%; Score 11; DB 2; Length 1009;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 574 VIYGDTSLEFV 584
|||||

DB 737 VIYGDTSLEFV 747
|||||

RESULT 9

A69312
DNA polymerase B1 (polB) homolog - Archaeoglobus fulgidus

C/Species: Archaeoglobus fulgidus

C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999

C/Accession: A69312

R/Klenk, H.P.; Claydon, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson

, J.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.

Glodek, A.; Zhou, L.; Overbeek, R.; Cocayne, J.D.; Feldman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A/Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Arriach, P.; Kaine, B.P.; Sykes, S.

Smith, H.O.; Woese, C.R.; Venter, J.C.

A/Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae

A/Reference number: A69250; MUID:98049343; PMID:9389475

A/Accession: A69312

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA
A/Residues: 1-781 <TLB>

A/Cross-references: GB:AE001070; GB:AE000782; NID:G268993; PIDN:AA90741.1; PID:G265013

C/Superfamily: herpesvirus DNA-directed DNA polymerase

Query Match
Best Local Similarity 1.2%; Score 10; DB 2; Length 781;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 239 YDPDIIVGVN 248
|||||

DB 218 YDPDIIVGVN 227
|||||

RESULT 10

DJB2L
DNA-directed DNA polymerase (EC 2.7.7.7) - human herpesvirus 4 (strain B95-8)

C/Species: human herpesvirus 4, Epstein-Barr virus

C/Date: 25-Feb-1985 #sequence_revision 25-Feb-1985 #text_change 11-Jun-1999

C/Accession: A00713; S33054

R/Bankier, A.T.; Deininger, P.L.; Farrell, P.J.; Barrell, B.G.

Wol. Biol. Med. 1, 21-45, 1983

A/Title: Sequence analysis of the 17,166 bp EcoRI fragment C of B95-8 Epstein-Barr virus

A/Reference number: A93065; MUID:8503573; PMID:6092825

A/Accession: A00713

A/Molecule type: DNA
A/Residues: 1-1015 <BAN>

A/Cross-references: EMBL:V01555; NID:G59074; PIDN:CAA24805.1; PID:G1334913

R/Baker, R.; Bankier, A.T.; Biggin, M.D.; Deininger, P.L.; Farrell, P.J.; Gibson, T.J.; H

Nature 310, 207-211, 1984

A/Title: DNA sequence and expression of the B95-8 Epstein-Barr virus genome.

A:Reference number: A03794; MUID:84270667; PMID:6087149
A:Content: annotation; protein coding region
C:Species: Herpesvirus DNA-directed DNA polymerase
C:Keywords: DNA binding; DNA replication; nucleotidyltransferase
Query Match 1.2%; Score 10; DB 1; Length 1015;
Best Local Similarity 100.0%; Pred. No. 0.48;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 574 YIGDTSLEFV 583
751 YIGDTSLEFV 760

RESULT 11
WMA012
DNA-directed DNA polymerase (EC 2.7.7.7) - human adenovirus 2
C:Species: Mastadenovirus h2 (human adenovirus 2)
A:Note: host Homo sapiens (man)
C>Date: 05-Apr-1983 #sequence_revision 05-Apr-1983 #text_change 16-Feb-1997
Accession: A92351; A92352; A00711
Gingerman, T.R.; Sclafy, D.; Gelinas, R.E.; Bing-Dong, J.; Yen, C.E.; Kelly, M.M.; Bull
Blot. Chem. 257 13475-13491, 1982
A:Title: Nucleotide sequences from the adenovirus-2 genome.
A:Reference number: A92351; MUID:83056843; PMID:7142161
A:Accession: A92351
A:Molecule type: DNA
A:Residues: 1-1056 <GIN>
R:Abstract: P.; Akuejar, G.; Pettegerson, M.; Pettegerson, U.
U. Biol. Chem. 257 13492-13498, 1982
A:Title: DNA sequence analysis of the region encoding the terminal protein and the hypoc
A:Reference number: A92352; MUID:83056844; PMID:7142162
A:Accession: A92352
A:Molecule type: DNA
A:Residues: 1-1056 <ALE>
A:Superfamily: adenovirus DNA-directed DNA polymerase
C:Keywords: DNA binding; DNA replication; nucleotidyltransferase
Query Match 1.2%; Score 10; DB 1; Length 1056;
Best Local Similarity 100.0%; Pred. No. 0.49;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 576 YGDTSLFV 585
868 YGDTSLFV 877

RESULT 12
BADS1
DNA-directed DNA polymerase (EC 2.7.7.7) - human adenovirus 5
C:Species: Mastadenovirus h5 (human adenovirus 5)
A:Note: host Homo sapiens (man)
C>Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 11-Jun-1999
Accession: A00712
Decker, B.M.M.; van Ormondt, H.
Gene 27 115-120, 1984
A:Title: The nucleotide sequence of fragment HindIII-C of human adenovirus type 5 DNA (r
A:Reference number: A91508; MUID:84183604; PMID:6325298
A:Accession: A00712
A:Molecule type: DNA
A:Residues: 1-1056 <DEK>
A:Cross-references: EMBL:X02996; NID:G58484; PIDN:CA51882.1; PID:G58495
C:Superfamily: adenovirus DNA-directed DNA polymerase
C:Keywords: DNA binding; DNA replication; nucleotidyltransferase
Query Match 1.2%; Score 10; DB 1; Length 1056;
Best Local Similarity 100.0%; Pred. No. 0.49;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 576 YGDTSLFV 585
868 YGDTSLFV 877

RESULT 13
DUND12
DNA-directed DNA polymerase (EC 2.7.7.7) - human adenovirus 12
C:Species: Mastadenovirus h12 (human adenovirus 12)
C>Date: 30-Jun-1989 #sequence_revision 17-May-1996 #text_change 11-Jun-1999
Accession: S33933; A25770
R:Spiegel, J.
submitted to the EMBL Data Library, June 1993
A:Reference number: S33928
A:Accession: S33933
A:Molecule type: DNA
A:Residues: 1-1061 <SPR>
A:Cross-references: EMBL:X73487; NID:G31361; PIDN:CA51882.1; PID:G31367
R:Shu, L.; Hong, J.S.; Wei, Y.F.; Engler, J.A.
Gene 46 187-195, 1986
A:Title: Nucleotide sequence of the genes encoded in early region 2b of human adenovirus
A:Reference number: A91557; MUID:87106854; PMID:3803925
A:Accession: A25770
A:Molecule type: DNA

A:Residues: 9231, 'S', 33-163, 'L', 163-180, 'YN', 183-460, 'T', 462-574, 'F', 576-891, 'T', 893-102
C:Comment: This gene is located on the E2B region of the genome.
C:Superfamily: adenovirus DNA-directed DNA polymerase
C:Keywords: DNA binding; DNA replication; early protein; nucleotidyltransferase
Query Match 1.2%; Score 10; DB 1; Length 1061;
Best Local Similarity 100.0%; Pred. No. 0.5;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 576 YGDTSLFV 585
874 YGDTSLFV 883

RESULT 14
S4169
DNA polymerase - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C>Date: 20-May-1994 #sequence_revision 26-May-1995 #text_change 09-Jun-2000
Accession: S41649
R:White, J.H.; Kilbey, B.J.; de Vries, E.; Goman, M.; Alano, P.; Cheesman, S.; McAlleese,
Nucleic Acids Res 21 3643-3646, 1993
A:Title: The gene encoding DNA polymerase alpha from Plasmodium falciparum.
A:Reference number: S41649; MUID:93376482; PMID:8367280
A:Accession: S41649
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-1855 <WHI>
A:Cross-references: EMBL:L18785
C:Superfamily: Plasmodium falciparum DNA polymerase
Query Match 1.2%; Score 10; DB 2; Length 1855;
Best Local Similarity 100.0%; Pred. No. 0.81;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 572 LKVIYDTS 581
1237 LKVIYDTS 1246

RESULT 15
E87278
sensor histidine kinase [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
Accession: E87278
R:Netman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eise, U.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.U.; Durkin, A.S.; Gilm, M.L.; Hatt, D.H.; Koon
n, J.; Esmolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venturi, J.C.; Frazer, C.M
Proc. Natl. Acad. Sci. U.S.A. 98 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11255647
A:Accession: E87278

A/Status: Preliminary
A/Molecule type: DNA
A/Residues: 1-514 <STO>
A/Cross-references: GB:AE005673; NID:g13421367; PIDN:AAK22225.1; GSPDB:GN00148
C/Genetics:
A/Gene: CC0238

Query Match
Best Local Similarity 1.1%; Score 9; DB 2; Length 534;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 27 IAENGERV 35
Db 181 IAENGERV 189

RESULT 16

75407
A/Status: probable DNA-directed DNA polymerase (EC 2.7.7.7) - Sulfolobus solfataricus
A/Alternate names: protein c04041
C/Species: Sulfolobus solfataricus
C/Date: 11-Oct-1997 #sequence_revision 24-Oct-1997 #text_change 20-Jun-2000
C/Accession: S75407
R/Sensen, C.W.; Klenk, H.P.; Singh, R.K.; Allard, G.; Chan, C.C.Y.; Liu, Q.Y.; Penny, S.
Mol. Microbiol. 22, 175-191, 1996
A/Title: Organizational characteristics and information content of an archaeal genome: 1
A/Reference number: S73076; MUID:97055432; PMID:8899719
A/Accession: S75407
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-764 <SEN>
A/Cross-references: EMBL:Y08257; NID:g1707772; PID:g1707813
A/Experimental source: strain P2
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1996.
C/Superfamily: herpesvirus DNA-directed DNA polymerase
C/Keywords: nucleotidyltransferase

Query Match
Best Local Similarity 1.1%; Score 9; DB 2; Length 764;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 728 KIGYIVKG 736
Db 691 KIGYIVKG 699

RESULT 17

5855
A/Status: DNA-directed DNA polymerase (EC 2.7.7.7) - Choristoneura biennis poxvirus
C/Species: Choristoneura biennis poxvirus
C/Date: 08-Jun-1994 #sequence_revision 10-Nov-1995 #text_change 08-Oct-1999
C/Accession: S25855
R/Mustafa, A.; Yuen, L.
DNA Seq. 2, 39-45, 1991
A/Title: Identification and sequencing of the Choristoneura biennis entomopoxvirus DNA
A/Reference number: S25855; MUID:92199242; PMID:1840516
A/Accession: S25855
A/Molecule type: DNA
A/Residues: 1-964 <MUS>
A/Cross-references: EMBL:X57314; NID:958845; PIDN:CAA40566.1; PID:G58846
C/Superfamily: herpesvirus DNA-directed DNA polymerase
C/Keywords: DNA binding; nucleotidyltransferase

Query Match
Best Local Similarity 1.1%; Score 9; DB 2; Length 964;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 575 IYGDTSLF 583
Db 750 IYGDTSLF 758

RESULT 18

JC5757

DNA-directed DNA polymerase (EC 2.7.7.7) III - yeast (Candida albicans)
A/Alternate names: DNA polymerase III
C/Species: Candida albicans
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 08-Dec-2000
C/Accession: JC5757; S60677
R/Nolan, T.; Rosamond, J.
Gene 183, 159-165, 1996
A/Title: Isolation and molecular characterization of the POL3 gene from Candida albicans
A/Reference number: JC5757; MUID:97149294; PMID:8996102
A/Accession: JC5757
A/Molecule type: DNA
A/Residues: 1-1038 <NOL>
A/Cross-references: EMBL:X88804
A/Note: Submitted to the EMBL Data Library, June 1995
C/Comment: This enzyme plays a role in DNA replication.
C/Genetics:
A/Gene: POL3
C/Superfamily: herpesvirus DNA-directed DNA polymerase
C/Keywords: nucleotidyltransferase; zinc finger
F/942-961/Region: zinc finger CCCC motif
F/992-1010/Region: zinc finger CCCC motif

Query Match
Best Local Similarity 1.1%; Score 9; DB 1; Length 1038;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 573 KVIYGDTS 581
Db 685 KVIYGDTS 693

RESULT 19

T18222
A/Status: DNA polymerase delta chain - yeast (Candida albicans)
C/Species: Candida albicans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Jun-2000
C/Accession: T18222
R/Bartell, B.G.; Rajandream, M.A.
Submitted to the EMBL Data Library, November 1998
A/Reference number: Z18831
A/Accession: T18222
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1038 <BAR>
A/Cross-references: EMBL:AL033396; PIDN:CAA21949.1
C/Genetics:
A/Note: Ca35A5.06C
C/Superfamily: herpesvirus DNA-directed DNA polymerase

Query Match
Best Local Similarity 1.1%; Score 9; DB 2; Length 1038;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 573 KVIYGDTS 581
Db 685 KVIYGDTS 693

RESULT 20

PC4394
A/Status: DNA-directed DNA polymerase (EC 2.7.7.7) - Ovine adenovirus OAV287 (fragment)
C/Species: Ovine adenovirus OAV287
C/Date: 10-Nov-1997 #sequence_revision 10-Nov-1997 #text_change 03-Nov-2000
C/Accession: PC4394
R/Vrati, S.; Brookes, D.E.; Boyle, D.B.; Both, G.W.
Gene 177, 35-41, 1996

A/Title: Nucleotide sequence of ovine adenovirus tripartite leader sequence and homolog
A/Reference number: JC5648; MUID:97080497; PMID:8921842
A/Accession: PC4394
A/Molecule type: DNA
A/Residues: 1-1080 <VRA>
A/Cross-references: GB:U11557; NID:g1117828; PIDN:AA55957.1; PID:g1117820

C/Comment: This enzyme is targeted to the nucleus by interaction with the terminal prote
C/Superfamily: adenovirus DNA-directed DNA polymerase
C/Keywords: nucleotidyltransferase

Query Match 1.1%; Score 9; DB 2; Length 1080;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 9; Conservative 0; Mismatches 0; Gaps 0;
Indels 0;
QY 575 XYGDTSLF 583
DB 882 XYGDTSLF 890

RESULT 21

hypothetical protein F10C2.4 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C/Accession: T20698
R/White, S.

Submitted to the EMBL Data Library, November 1996
Reference number: Z19311
Accession: T20698

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-1081 <N1L>
A/Cross-references: EMBL:Z81497; PIDN:CA804077.1; GSPDB:GN00023; CESP:F10C2.4

A/Experimental source: clone F10C2

C/Genetics:

A/Gene: CESP:F10C2.4

A/Map position: 5

A/Intons: 29/3; 112/3; 216/3; 281/2; 524/3; 862/3; 1012/2

C/Superfamily: herpesvirus DNA-directed DNA polymerase

QY 573 XYGDTSLF 581

DB 721 XYGDTSLF 729

RESULT 22

DJB228

DNA-directed DNA polymerase (EC 2.7.7.7) - human herpesvirus 3
C/Species: human herpesvirus 3, varicella-zoster virus
C/Date: 30-Sep-1988 #sequence_revision 30-Sep-1988 #text_change 11-Jun-1999

C/Accession: B27214

A/Title: The complete DNA sequence of varicella-zoster virus.

A/Reference number: A27345; M01D:86306657; PMID:3018124

A/Accession: B27214

A/Molecule type: DNA

A/Residues: 1-1194 <DAN>

A/Cross-references: EMBL:X04370; NID:G59989; PIDN:CAA27911.1; PID:G60017

C/Genetics:

A/Gene: 28

C/Superfamily: herpesvirus DNA-directed DNA polymerase

C/Keywords: DNA binding; DNA replication; nucleotidyltransferase

Query Match 1.1%; Score 9; DB 1; Length 1194;
Best Local Similarity 100.0%; Pred. No. 5.6;
Matches 9; Conservative 0; Mismatches 0; Gaps 0;
Indels 0;

QY 573 XYGDTSLF 581

DB 846 XYGDTSLF 854

RESULT 23

A33602

DNA-directed DNA polymerase (EC 2.7.7.7) REV3 - yeast (Saccharomyces cerevisiae)
N/Alternate names: protein p255; protein YPL167C
C/Species: Saccharomyces cerevisiae
C/Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 20-Jun-2000
C/Accession: A33602; S65178; S69432
R/Morrison, A.; Christensen, R.B.; Alley, J.; Beck, A.K.; Bernatine, E.G.; Lemont, J.F
J. Bacteriol. 171, 5659-5667, 1989
A/Title: REV3, a Saccharomyces cerevisiae gene whose function is required for induced m
A/Reference number: A33602; M01D:90008808; PMID:2676986
A/Accession: A33602
A/Molecule type: DNA
A/Residues: 1-1504 <NOR>
A/Cross-references: GB:M29683; NID:G172386; PIDN:AAA4968.1; PID:G172387
A/Experimental source: strain S288C
R/Purnelle, B.; Coester, F.; Goffeau, A.
Submitted to the Protein Sequence Database, May 1996
A/Reference number: S65154
A/Accession: S65178
A/Molecule type: DNA
A/Residues: 1-1504 <PUM>
A/Cross-references: EMBL:Z73523; NID:G1370352; PIDN:CAA97873.1; PID:G1370353; MIPS:YPL167C
A/Experimental source: strain S288C (A5972)
R/Purnelle, B.; Comblez, S.; Coester, F.; Naveau, F.; Goffeau, A.
Submitted to the EMBL Data Library, March 1996
A/Description: The sequence of 55 kb on the left arm of yeast chromosome XVI identifies
A/Reference number: S69428
A/Accession: S69432
A/Molecule type: DNA
A/Residues: 1-1504 <PRM>
A/Cross-references: EMBL:X96770; NID:G1403537; PIDN:CAA6554.1; PID:G1403542
C/Genetics:

A/Gene: SGD:REV3; PS01

A/Cross-references: SGD:S0006088; MIPS:YPL167C

A/Map position: 16L

C/Superfamily: yeast DNA-directed DNA polymerase REV3

C/Keywords: DNA binding; nucleotidyltransferase

Query Match 1.1%; Score 9; DB 2; Length 1504;
Best Local Similarity 100.0%; Pred. No. 6.9;
Matches 9; Conservative 0; Mismatches 0; Gaps 0;
Indels 0;

QY 576 YGDTSLFV 584

DB 1140 YGDTSLFV 1148

RESULT 24

DJB228

DNA-directed DNA polymerase (EC 2.7.7.7) delta catalytic chain - slime mold (Physarum pc

C/Species: Physarum polycephalum

C/Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 18-Jun-1999

C/Accession: S22088; S65029

R/Holler, E.

Submitted to the EMBL Data Library, May 1992

A/Reference number: S22088

A/Accession: S22088

A/Molecule type: mRNA

A/Residues: 1-148 <HOL>

A/Cross-references: EMBL:X66049; NID:G3192; PIDN:CAA46847.1; PID:G3193

A/Experimental source: strain M3C VIII

R/Achhammer, G.; Winkler, A.; Angerer, B.; Holler, E.

Cut. Genet. 28, 534-545, 1995

A/Title: DNA polymerase delta of Physarum polycephalum.

A/Reference number: S65029; M01D:96132109; PMID:8593684

A/Accession: S65029

A/Molecule type: mRNA

A/Residues: 1-142 <ACH>

A/Cross-references: EMBL:X66049

C/Genetics:

A/Gene: POL3

C/Superfamily: herpesvirus DNA-directed DNA polymerase

C/Keywords: DNA replication; nucleotidyltransferase

Query Match
Best Local Similarity 1.0%; Score 8; DB 2; Length 148;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 574 VIYGDIDS 581
Db 141 VIYGDIDS 148

RESULT 25

T20848

hypothetical protein F13E9.10 - *Caenorhabditis elegans*C/Species: *Caenorhabditis elegans*

C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000

C/Accession: T20848

R/McMurry, A. Submitted to the EMBL Data Library, February 1996

A/Accession: T20848

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-229 <WIL>

A/Cross-references: EMBL:Z69383; PIDN:CAA93413.1; GSPDB:GN00022; CESP:F13E9.10

A/Experimental source: clone F13E9

C/Genetics:

A/Map position: 4

A/Intons: 65/3; 127/3

C/Superfamily: *Caenorhabditis elegans* hypothetical protein F13E9.10

Query Match
Best Local Similarity 1.0%; Score 8; DB 2; Length 223;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 700 LEEYTTTA 707
Db 165 LEEYTTTA 172

RESULT 26

A99284

hypothetical protein SSO1290 [imported] - *Sulfolobus solfataricus*C/Species: *Sulfolobus solfataricus*

C/Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001

C/Accession: A99284

She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aways, M.J.; Chan-

drett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.

Submitted to Genbank, April 2001

A/Description: *Sulfolobus solfataricus* complete genome.

A/Reference number: A99139

A/Accession: A99284

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-324 <KUR>

A/Cross-references: GB:AE006641; NID:913814490; PIDN:AAK41528.1; GSPDB:GN00155

C/Genetics:

A/Map position: 2

A/Intons: 105/2; 272/2; 394/3

A/Accession: T33878

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-414 <ROH>

A/Cross-references: EMBL:AF125447; PIDN:AD12807.1; GSPDB:GN00020; CESP:Y14H12B.2

A/Experimental source: strain Bristol NZ; clone Y14H12B

C/Genetics:

A/Map position: 2

A/Intons: 105/2; 272/2; 394/3

A/Accession: T33878

A/Status: preliminary

A/Molecule type: DNA

A/Note: Nostoc sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120

C/Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002

C/Accession: AB1828

R/Xanezo, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi

Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S

DNA Res. 8, 205-213, 2001

A/Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Ana*

A/Reference number: AB1807; MID:21595285; PMID:1175840

A/Accession: AB1828

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-365 <KUR>

A/Cross-references: GB:BA000019; PIDN:BA07694.1; PID:917135148; GSPDB:GN00179

A/Experimental source: strain PCC 7120

C/Genetics:

A/Map position: 4

A/Intons: 65/3; 127/3

A/Accession: T33878

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-414 <ROH>

A/Cross-references: EMBL:AF125447; PIDN:AD12807.1; GSPDB:GN00020; CESP:Y14H12B.2

A/Experimental source: strain Bristol NZ; clone Y14H12B

C/Genetics:

A/Map position: 2

A/Intons: 105/2; 272/2; 394/3

A/Accession: T33878

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-414 <ROH>

A/Cross-references: EMBL:AF125447; PIDN:AD12807.1; GSPDB:GN00020; CESP:Y14H12B.2

A/Experimental source: strain Bristol NZ; clone Y14H12B

C/Genetics:

A/Map position: 2

A/Intons: 105/2; 272/2; 394/3

A/Accession: T33878

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-414 <ROH>

A/Cross-references: EMBL:AF125447; PIDN:AD12807.1; GSPDB:GN00020; CESP:Y14H12B.2

A/Experimental source: strain Bristol NZ; clone Y14H12B

C/Genetics:

A/Map position: 2

A/Intons: 105/2; 272/2; 394/3

A/Accession: T33878

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-414 <ROH>

A/Cross-references: EMBL:AF125447; PIDN:AD12807.1; GSPDB:GN00020; CESP:Y14H12B.2

A/Experimental source: strain Bristol NZ; clone Y14H12B

C/Genetics:

A/Map position: 2

A/Intons: 105/2; 272/2; 394/3

A/Accession: T33878

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-414 <ROH>

A/Cross-references: EMBL:AF125447; PIDN:AD12807.1; GSPDB:GN00020; CESP:Y14H12B.2

A/Experimental source: strain Bristol NZ; clone Y14H12B

C/Genetics:

A/Map position: 2

A/Intons: 105/2; 272/2; 394/3

A/Accession: T33878

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-414 <ROH>

A/Cross-references: EMBL:AF125447; PIDN:AD12807.1; GSPDB:GN00020; CESP:Y14H12B.2

A/Experimental source: strain Bristol NZ; clone Y14H12B

C/Genetics:

A/Map position: 2

A/Intons: 105/2; 272/2; 394/3

A/Accession: T33878

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-419 <STO>
 A/Cross-references: GB:AL450380; NID:G13092747; PIDN:CAC30051.1; GSPDB:GN00147
 C/Genetics:
 A/Gene: dfrp
 C/Superfamily: pantothenate metabolism flavoprotein dfr

Query Match 1.0%; Score 8; DB 2; Length 419;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 419 RGATVLEP 426
 DB 140 RGATVLEP 147

RESULT 30
 D72383
 NADH oxidase - Thermotoga maritima (strain MSB8)
 C/Species: Thermotoga maritima
 C/Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
 C/Accession: D72383

Neilsen, K.E.; Clayton, R.A.; Gili, S.R.; Gwin, M.L.; Dodson, R.U.; Haft, D.H.; Hickey
 Bartlett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.
 C.M.
 Nature 399, 323-329, 1999
 A/Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
 A/Reference number: A72200; MUID:99287316; PMID:10360571
 A/Accession: D72383
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-443 <ARN>
 A/Cross-references: GB:AE001718; GB:AE000512; NID:G4980881; PIDN:ADJ35465.1; PID:G498088
 C/Genetics:
 A/Experimental source: strain MSB8
 A/Gene: TM0379
 C/Superfamily: NADH peroxidase; dihydroliipoamide dehydrogenase homology
 F/e-429/Domain: dihydroliipoamide dehydrogenase homology <DUD>

Query Match 1.0%; Score 8; DB 2; Length 443;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 567 ARKGLKRV 574
 DB 245 ARKGLKRV 252

RESULT 31

E82675

Heat shock protein Xf1485 [imported] - Xylella fastidiosa (strain 9a5c)
 C/Species: Xylella fastidiosa
 C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 17-Nov-2000
 C/Accession: A82675

Ranounous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
 Nature 406, 151-157, 2000
 A/Title: The genome sequence of the plant pathogen Xylella fastidiosa.
 A/Reference number: A82515; MUID:20365717; PMID:10910347
 A/Note: for a complete list of authors see reference number A59328 below
 A/Accession: A82675
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-459 <SIM>
 A/Cross-references: GB:AE003978; GB:AE003849; NID:G9106510; PIDN:AMF84294.1; GSPDB:GN001

A/Experimental source: strain 9a5c
 R/Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
 Brites, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carter, H
 as-Melo, E.; Docena, C.; El-Dorzy, H.; Facinanti, A.P.; Ferreira, A.J.S.
 submitted to GenBank, June 2000
 A/Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; From
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laizy
 Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
 A/Authors: Martins, E.M.F.; Matukuma, A.Y.; Menck, C.F.M.; Miranda, E.C.; Miyaki, C.Y.
 F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmeri, D.A

Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Sanello, R.V.; Sawasgi,
 A/Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
 M.; Tushneto, M.H.; Valada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.;
 A/Reference number: A59328
 A/Contents: annotation
 C/Genetics:
 A/Gene: Xf1485
 C/Superfamily: heat shock protein hsp70; FtsH/SEC18/CDC48-type ATP-binding domain homolo

Query Match 1.0%; Score 8; DB 2; Length 459;
 Best Local Similarity 100.0%; Pred. No. 25;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 489 TVLERLLR 496
 DB 413 TVLERLLR 420

RESULT 32

E83827

hypothetical protein BH1421 [imported] - Bacillus halodurans (strain C-125)
 C/Species: Bacillus halodurans
 C/Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
 C/Accession: E83827

R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hir
 Nucleic Acids Res 28, 4317-4331, 2000
 A/Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
 A/Reference number: A83650; MUID:20512582; PMID:11058132
 A/Accession: E83827
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-514 <STO>
 A/Cross-references: GB:AP001512; GB:BA000004; NID:G10174030; PIDN:BA05140.1; GSPDB:GN0C
 A/Experimental source: strain C-125
 C/Genetics:
 A/Gene: BH1421

Query Match 1.0%; Score 8; DB 2; Length 514;
 Best Local Similarity 100.0%; Pred. No. 28;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 518 LDEROKAL 525
 DB 482 LDEROKAL 489

RESULT 33

C81273

hypothetical protein Cj1302 [imported] - Campylobacter jejuni (strain NCTC 11168)
 C/Species: Campylobacter jejuni
 C/Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002
 C/Accession: C81273

R/Parkhill, J.; Wren, B.W.; Wungall, K.; Kelsey, J.M.; Churcher, C.; Baeham, D.; Chiller
 C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Barril
 Nature 403, 665-668, 2000
 A/Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals h
 A/Reference number: A81250; MUID:20150912; PMID:10688204
 A/Accession: C81273
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-520 <PAR>
 A/Cross-references: GB:AL139078; GB:AL111168; NID:G6968723; PIDN:CAB73729.1; PID:G696873
 A/Experimental source: serotype O2, strain NCTC 11168
 C/Genetics:
 A/Gene: Cj1302

Query Match 1.0%; Score 8; DB 2; Length 520;
 Best Local Similarity 100.0%; Pred. No. 28;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 492 ERLLELRK 499
 DB 97 ERLLELRK 104

RESULT 34

S49439

glucose-1-phosphate adenylyltransferase (EC 2.7.7.27) - maize

C/Species: Zea mays (maize)

C/Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 18-Jun-1999

C/Accession: S49439

R/Smith-White, B.J.; Gilroy, M.J.; Gilmore, V.; Preiss, J.; Hannah, L.C.

A/Description: The EMBL Data Library, October 1994

A/Reference number: S49439

A/Accession: S49439

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1521 <SMI>

A/Cross-references: EMBL:Z28111; NID:G558364; PIDN:CAA86227.1; PID:G558365

A/Function:

A/Description: catalyzes the formation of ADPGlucose and pyrophosphate from alpha-D-glucose-1-phosphate and ATP

C/Suprafamily: glucose-1-phosphate adenylyltransferase

C/Keywords: glycogen/starch biosynthesis; nucleotidyltransferase

Query Match

Best Local Similarity 1.0%; Score 8; DB 2; Length 521;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 453 GPDTPVP 460

DB 66 GPDTPVP 73

RESULT 35

E87086

probable regulatory protein [imported] - Mycobacterium leprae

C/Species: Mycobacterium leprae

C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001

C/Accession: E87086

R/Cole, S.T.; Eiglsmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; He

R.; Davies, R.M.; Devlin, K.; Duthey, S.; Feltwell, T.; Frazer, A.; Hamlin, N.; Holroyd,

Gamm, M.A.; Rutherford, K.M.

A/Authors: Rutherford, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Sq

A/Title: Massive gene decay in the leprosy bacillus.

A/Reference number: A86909; MUID:21128732; PMID:11234002

A/Accession: E87086

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-563 <STO>

A/Cross-references: GB:AL450380; NID:G13093298; PIDN:GAC30370.1; GSPDB:GN00147

C/Genetics:

A/Gene: M1419

Query Match

Best Local Similarity 1.0%; Score 8; DB 2; Length 563;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 768 PAALRIG 775

DB 286 PAALRIG 293

RESULT 36

C89876

hypothetical protein purL [imported] - Staphylococcus aureus (strain N315)

C/Species: Staphylococcus aureus

C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001

C/Accession: C89876

R/Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguni

ma, A.; Mizutani, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;

C.; Shibata, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.

A/Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.

A/Reference number: A89758; MUID:21111952; PMID:11418146

A/Accession: C89876

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-729 <KUR>

A/Cross-references: GB:BA000018; PID:G13700870; PIDN:BA02166.1; GSPDB:GN00149

A/Experimental source: strain N315

C/Genetics:

A/Suprafamily: phosphoribosylformylglycinamide synthase component II

Query Match

Best Local Similarity 1.0%; Score 8; DB 2; Length 729;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 605 EIKLEKRY 612

DB 11 EIKLEKRY 18

RESULT 37

JDBC22

DNA-directed DNA polymerase (EC 2.7.7.7) II - Escherichia coli (strain K-12)

C/Species: Escherichia coli

C/Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 01-Mar-2002

C/Accession: S15943; S19262; S40576; J00780; A38840; S19263; S12820; B36236; D64727; S54

R/Iwaaki, H.; Ishino, Y.; Toh, H.; Nakata, A.; Shimagawa, H.

Mol. Gen. Genet. 226, 24-33, 1991

A/Title: Escherichia coli DNA polymerase II is homologous to alpha-like DNA polymerases

A/Reference number: S15943; MUID:91238699; PMID:2034216

A/Accession: S15943

A/Molecule type: DNA

A/Residues: 1-783 <IWA>

A/Cross-references: EMBL:X54847; NID:G42462; PIDN:CAA38616.1; PID:G581193

A/Experimental source: strain K-12, substrain W3110

A/Accession: S19262

A/Molecule type: protein

A/Residues: 2-9 <IWA2>

R/Tura, T.; Mori, H.; Nagai, H.; Nagata, T.; Ishihama, A.; Fujita, N.; Isono, K.; Mizobu

submitted to the EMBL Data Library, December 1992

A/Description: Systematic sequencing of the Escherichia coli genome: analysis of the 0-2

A/Reference number: S40531

A/Accession: S40576

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 'V', 2-783 <YUR>

A/Cross-references: EMBL:D10483; NID:G216434

A/Experimental source: strain K-12

R/Bonner, C.A.; Hays, S.; McIntee, K.; Goodman, M.F.

Proc. Natl. Acad. Sci. U.S.A. 87, 7663-7667, 1990

A/Title: DNA polymerase II is encoded by the DNA damage-inducible dna gene of Escherich

A/Reference number: J00780; MUID:91017565; PMID:2217198

A/Accession: J00780

A/Molecule type: DNA

A/Residues: 1-256 'DD', 259-271 'G', 273-458 <BON>

A/Cross-references: GB:M37727; NID:G145744; PIDN:AAA23684.1; PID:G145746

A/Accession: A38840

A/Molecule type: protein

A/Residues: 2-9, 'XQ', 12-21, 'H', 23, 'X', 25-28 <BO2>

R/Chen, H.

submitted to the EMBL Data Library, June 1990

A/Reference number: S19263

A/Accession: S19263

A/Molecule type: DNA

A/Residues: 1-171, 'A', 173-734, 'T', 736-739, 'PGPFTTGLRTSDPPATRRG', 763, 'NT', 766-767, 'Y'

A/Cross-references: EMBL:M35311; NID:G147319; PIDN:AAA24407.1; PID:G147320

Nucleic Acids Res. 18, 7185-7186, 1990

A:Residues: 90-171, 'A', 173-734, 'T', 736-739, 'PGLPTTGLRLTSPDPRATRG', 763, 'NT', 766-767, 'A'
A:Cross-references: EMBL:M35371
R:Chen, H.; Sun, Y.; Stark, T.; Beattie, W.; Moses, R.E.
DNA Cell Biol. 9, 631-635, 1990
A:Title: Nucleotide sequence and deletion analysis of the polB gene of *Escherichia coli*.
A:Reference number: A36236; MUID:91083835; PMID:2261080
A:Accession: B36236
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-783 <BLAT>
A:Cross-references: GB:BA000116; GB:U00096; NID:G1786240; PIND:AAC73171.1; PID:G1786246;
Experimental source: strain K-12, substrain MG1655
A:Accession: S54807
A:Reference number: S54807
A:Accession: S54807
A:Molecule type: mRNA
A:Residues: 428-444, 'LIQSTVPRV', 454, 'SMPGSRKINAY', 467-482, 'RVTV', 487-510, 'L', 512-520, 'T'
A:Cross-references: EMBL:A87158; NID:9809066; PID:9809067
A:Note: The source was identified as *Physarum polycephalum*, this is probably a misassignment
A:Note: compared to *Escherichia coli* there is 96% sequence identity at the DNA level and
R:Ackhammer, G.; Winkler, A.; Angerer, B.; Holler, E.
Curr. Genet. 28, 534-545, 1995
A:Title: DNA polymerase delta of *Physarum polycephalum*.
A:Reference number: S65029; MUID:96132109; PMID:8593684
A:Accession: S65030
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: mRNA
A:Residues: 428-444, 'LIQSTVPRV', 454, 'SMPGSRKINAY', 467-482, 'RVTV', 487-510, 'L', 512-520, 'T'
A:Cross-references: EMBL:A87158
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1995
A:Note: the source was identified as *Physarum polycephalum*; this is probably a misassignment
A:Note: compared to *Escherichia coli* there is 96% sequence identity at the DNA level and
C:Genetics:
A:Gene: polB
A:Map position: 2 min
A:Start codon: GTG
A:Function:
A:Pathway: purine metabolism; pyrimidine metabolism
A:Superfamily: *Escherichia coli* DNA-directed DNA polymerase II
A:Keywords: DNA binding; DNA biosynthesis; nucleic acid metabolism
A:Product: DNA-directed DNA polymerase II [status experimental] <MAT>
Query Match 1.0%; Score 8; DB 1; Length 783;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 574 VIYGDTS 581
Db 541 VIYGDTS 548
RESULT 38
DNA polymerase II [imported] - *Escherichia coli* (strain O157:H7, substrain RMD 0509952)
C:Species: *Escherichia coli*
C:Date: 16-Jul-2001 #sequence_revision 16-Jul-2001 #ext_change 03-Aug-2001
C:Accession: H90636
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
Gawara, N.; Yasunaga, T.; Kuwara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and gene
A:Reference number: A99629; MUID:21156231; PMID:11258796

A:Accession: H90636
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-783 <HAV>
A:Cross-references: GB:BA000007; PIND:BA33467.1; PID:G13359520; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RMD 0509952
C:Genetics:
A:Gene: ECG064
C:Superfamily: *Escherichia coli* DNA-directed DNA polymerase II
Query Match 1.0%; Score 8; DB 2; Length 783;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 574 VIYGDTS 581
Db 541 VIYGDTS 548
RESULT 39
DNA polymerase II [imported] - *Escherichia coli* (strain O157:H7, substrain EDL933)
C:Species: *Escherichia coli*
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #ext_change 14-Sep-2001
C:Accession: H85487
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Hiller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
A:Reference number: A85480; MUID:21074535; PMID:11206551
A:Accession: H85487
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-783 <STO>
A:Cross-references: GB:BA005174; NID:G12512757; PIND:AAG54364.1; GSPDB:GN00145; UNCP:200
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: polB
A:Superfamily: *Escherichia coli* DNA-directed DNA polymerase II
Query Match 1.0%; Score 8; DB 2; Length 783;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 574 VIYGDTS 581
Db 541 VIYGDTS 548
RESULT 40
DNA polymerase II [imported] - *Salmonella enterica* subsp. *enterica* serovar Typh1 (strain
A:Species: *Salmonella enterica* subsp. *enterica* serovar Typh1
A:Note: this species has also been called *Salmonella typh1*
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #ext_change 27-Nov-2001
C:Accession: AB0514
R:Parthill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
Th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Doud, L.; White, N.; Farrar,
S.; Moulé, S.; O'Garra, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serov.
A:Reference number: AB0502; PMID:11677608
A:Accession: AB0514
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-783 <PAR>
A:Cross-references: GB:AL513382; PIND:CAD01253.1; PID:G16501382; GSPDB:GN00176
C:Genetics:
A:Gene: STY0112
C:Superfamily: *Escherichia coli* DNA-directed DNA polymerase II
Query Match 1.0%; Score 8; DB 2; Length 783;

Best Local Similarity 100.0%; Pred. No. 40;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 574 VIYGDTS 581
DB 541 VIYGDTS 548

RESULT 41

DNA polymerase II PA186 [imported] - Pseudomonas aeruginosa (strain PA01)

C/Species: Pseudomonas aeruginosa

C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

C/Accession: G83410

R/Stover, C.K.; Pham, X.O.; Errin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B.

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,

Loay, S.; Olson, M.V.

Nature 406, 959-964, 2000

Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path

Accession: G83410

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-787 <STO>

A/Cross-references: GB:AB004614; GB:AE004091; NID:G947868; PIDN:AG05275.1; GSPDB:GN001

A/Experimental source: strain PA01

C/Genetics:

A/Gene: polB; PA186

C/Superfamily: Escherichia coli DNA-directed DNA polymerase II

Query Match 1.0%; Score 8; DB 2; Length 787;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 574 VIYGDTS 581
DB 543 VIYGDTS 550

RESULT 42

DNA polymerase II VC1212 [imported] - Vibrio cholerae (strain N16961 serogroup O1)

C/Species: Vibrio cholerae

C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001

C/Accession: E82227

R/Haidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gilm, M.L.; Dodson, R.J.;

Hardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, F.

R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A/Reference number: A82035; MUID:20406833; PMID:10952301

A/Accession: E82227

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-787 <HEI>

A/Cross-references: GB:AB004201; GB:AE003852; NID:G9555689; PIDN:AAE94371.1; GSPDB:GN001

A/Experimental source: serogroup O1; strain N16961; biotype El Tor

C/Genetics:

A/Gene: VC1212

A/Map position: 1

C/Superfamily: Escherichia coli DNA-directed DNA polymerase II

Query Match 1.0%; Score 8; DB 2; Length 787;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 574 VIYGDTS 581
DB 544 VIYGDTS 551

RESULT 43

AE0064

DNA-directed DNA polymerase (EC 2.7.7.7) [imported] - Yersinia pestis (strain CO92)

C/Species: Yersinia pestis

C/Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 27-Nov-2001

C/Accession: AE0064

R/Parhill, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M.B.

deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;

Il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,

Nature 413, 523-527, 2001

Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A/Reference number: AB0001; MUID:21470413; PMID:11586360

A/Accession: AE0064

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-789 <KUP>

A/Cross-references: GB:AL590842; PIDN:CA089376.1; PID:G15978613; GSPDB:GN00175

C/Genetics:

A/Gene: YPO0518

C/Superfamily: Escherichia coli DNA-directed DNA polymerase II

C/Keywords: nucleotidyltransferase

Query Match 1.0%; Score 8; DB 2; Length 789;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 574 VIYGDTS 581
DB 546 VIYGDTS 553

RESULT 44

DNA-directed DNA polymerase (EC 2.7.7.7) B - Cenarchaeum symbiosum

C/Species: Cenarchaeum symbiosum

C/Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 18-Feb-2000

C/Accession: T31321

R/Schleper, C.; DeLong, E.F.; Preston, C.M.; Feldman, R.A.; Wu, K.Y.; Swanson, R.V.

J. Bacteriol. 180, 5003-5009, 1998

Title: Genomic analysis reveals chromosomal variation in natural populations of the ur

A/Reference number: Z20994; MUID:98422450; PMID:9748430

A/Accession: T31321

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-844 <SCH>

A/Cross-references: EMBL:AF083072; NID:G3599393; PID:G3599407; PIDN:AA062712.1

C/Genetics:

A/Gene: pol1

C/Superfamily: herpesvirus DNA-directed DNA polymerase

C/Keywords: nucleotidyltransferase

Query Match 1.0%; Score 8; DB 2; Length 844;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 576 YGPTDSL 583
DB 619 YGPTDSL 626

RESULT 45

DNA-directed DNA polymerase (EC 2.7.7.7) I - Pyrodicticum occultum

C/Species: Pyrodicticum occultum

C/Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 16-Feb-1997

C/Accession: A56277

R/Uemori, T.; Ishino, Y.; Doi, H.; Kato, I.

J. Bacteriol. 177, 2164-2177, 1995

Title: The hyperthermophilic archaeon Pyrodicticum occultum has two alpha-like DNA poly

A/Reference number: A56277; MUID:95238290; PMID:7721707

A/Accession: A56277

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-879 <UEM>

A/Cross-references: GB:D38573

C:Genetics:
A:Gene: polA
C:Superfamily: herpesvirus DNA-directed DNA polymerase
C:Keywords: nucleotidyltransferase

Query Match
Best Local Similarity 100.0%; Score 8; DB 2; Length 879;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 576 YGDTDSL 583
DB 654 YGDTDSL 661

RESULT 46

S23019

DNA-directed DNA polymerase (EC 2.7.7.7) - Sulfolobus solfataricus

C:Species: Sulfolobus solfataricus

C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 18-Jun-1999

C:Accession: S23019

C:Keywords: F.M.; de Martino, C.; Rossi, M.

C:Title: A DNA polymerase from the archaeon Sulfolobus solfataricus shows sequence simi

A:Reference number: S23019; MUID:92310966; PMID:1614858

A:Accession: S23019

A:Molecule type: DNA

A:Residues: 1882 <PIS>

A:Cross-references: EMBL:X64466; NID:947564; PIDD:CAA5795.1; PID:947565

A:Experimental source: strain MT 4

C:Genetics:

A:Gene: polS

C:Superfamily: herpesvirus DNA-directed DNA polymerase

C:Keywords: DNA binding; DNA replication; nucleotidyltransferase

C:Query Match

Best Local Similarity 100.0%; Score 8; DB 2; Length 882;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 576 YGDTDSL 583

DB 653 YGDTDSL 660

RESULT 47
P90201

DNA polymerase I (DNA polymerase B) (dpol) [imported] - Sulfolobus solfataricus

C:Species: Sulfolobus solfataricus

C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 17-May-2002

C:Accession: P90201

C:Keywords: R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-

C:Title: A DNA polymerase from the archaeon Sulfolobus solfataricus shows sequence simi

A:Reference number: A99139

A:Accession: P90201

A:Molecule type: DNA

A:Residues: 1-882 <KUR>

A:Cross-references: GB:AEO0641; NID:913813715; PIDD:AAK40869.1; GSPDB:GN00155

A:Gene: dpol

C:Superfamily: herpesvirus DNA-directed DNA polymerase

C:Query Match

Best Local Similarity 100.0%; Score 8; DB 2; Length 882;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 576 YGDTDSL 583

DB 653 YGDTDSL 660

RESULT 48
T11742

egg sperm receptor - sea urchin (Strongylocentrotus purpuratus)

C:Species: Strongylocentrotus purpuratus (purple urchin)

C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 24-Nov-1999

C:Accession: T11742

C:Keywords: K.R.; Partin, J.S.; Lennarz, W.J.

C:Title: Sea urchin egg receptor for sperm: sequence similarity of binding domain and hs

A:Reference number: T11742; MUID:93197888; PMID:8383878

A:Accession: T11742

A:Molecule type: mRNA

A:Residues: 1-889 <FOL>

A:Cross-references: EMBL:L04969; NID:91580782; PID:91580783

A:Experimental source: immature ovary

C:Superfamily: heat shock protein 91

C:Query Match

Best Local Similarity 100.0%; Score 8; DB 2; Length 889;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 525 LKVLNAN 532

DB 220 LKVLNAN 227

RESULT 49
A42543

DNA-directed DNA polymerase (EC 2.7.7.7) - Chlorella virus PBCV-1

C:Species: Chlorella virus PBCV-1

C>Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999

C:Accession: A42543

C:Keywords: R.; Grabherr, R.; Straaser, P.; Van Etten, J.L.

C:Title: The DNA polymerase gene from Chlorella viruses PBCV-1 and NY-2A contains an int

A:Reference number: A42543; MUID:92263776; PMID:1585643

A:Accession: A42543

A:Molecule type: DNA

A:Residues: 1-913 <GRA>

A:Cross-references: GB:M86837; NID:9323321; PIDD:AAA88827.1; PID:9323322

C:Genetics:

C:Superfamily: herpesvirus DNA-directed DNA polymerase

C:Query Match

Best Local Similarity 100.0%; Score 8; DB 1; Length 913;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 574 VIYGDTS 581

DB 635 VIYGDTS 642

RESULT 50
B42543

DNA-directed DNA polymerase (EC 2.7.7.7) - Chlorella virus CV-NY-2A

C:Species: Chlorella virus CV-NY-2A

C>Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 11-Jun-1999

C:Accession: B42543

C:Keywords: R.; Straaser, P.; Van Etten, J.L.

C:Title: The DNA polymerase gene from Chlorella viruses PBCV-1 and NY-2A contains an int

A:Reference number: A42543; MUID:92263776; PMID:1585643

A:Accession: B42543

A:Molecule type: DNA

A:Residues: 1-913 <GRA>

A:Cross-references: GB:M86837; NID:9323321; PIDD:AAA88827.1; PID:9323322

C:Genetics:

C:Superfamily: herpesvirus DNA-directed DNA polymerase

C:Query Match

Best Local Similarity 100.0%; Score 8; DB 1; Length 913;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 574 VIYGDTS 581

DB 635 VIYGDTS 642

A:introns: 682/3
 C:Superfamily: herpesvirus DNA-directed DNA polymerase
 C:Keywords: DNA binding; DNA biosynthesis; DNA replication; exonuclease; nucleotidyltransferase
 F:182-199/Region: exonuclease pattern A
 F:268-282/Region: exonuclease pattern B
 F:388-401/Region: exonuclease pattern C

Query Match

1.0%; Score 8; DB 1; Length 913;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 574 VIYGDTDS 581
 |||||
 Db 635 VIYGDTDS 642

RESULT 51

1.0%; Score 8; DB 1; Length 913;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1.0%; Score 8; DB 1; Length 913;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1.0%; Score 8; DB 2; Length 913;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 52

1.0%; Score 8; DB 2; Length 913;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1.0%; Score 8; DB 2; Length 913;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1.0%; Score 8; DB 2; Length 913;
 Best Local Similarity 100.0%; Pred. No. 47;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 314 AFDIEVYS 321

RESULT 53

1.0%; Score 8; DB 1; Length 913;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1.0%; Score 8; DB 1; Length 913;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1.0%; Score 8; DB 1; Length 913;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1.0%; Score 8; DB 1; Length 913;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1.0%; Score 8; DB 1; Length 913;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 574 VIYGDTDS 581
 |||||
 Db 736 VIYGDTDS 743

RESULT 54

1.0%; Score 8; DB 1; Length 913;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1.0%; Score 8; DB 1; Length 913;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1.0%; Score 8; DB 1; Length 913;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 574 VIYGDTDS 581
 |||||
 Db 736 VIYGDTDS 743

RESULT 55

1.0%; Score 8; DB 1; Length 913;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Probable DNA-directed DNA polymerase (EC 2.7.7.7) U38 [similarity] - human herpesvirus 6
 C:Species: strain herpesvirus 6
 A:Virus: strain 229
 C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 05-May-2000
 C:Accession: T44185
 R:Dominy: G.; Dambach, T.R.; Stamey, F.R.; Dewhurst, S.; Inoue, N.; Pellett, P.E.
 J: Virol. 73, 8040-8052, 1999
 A:Title: Human herpesvirus 68 genome sequence: coding content and comparison with human
 A:Reference number: Z22734; MUID:9941318; PMID:10482553
 A:Accession: T44185
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1012 <DOM>
 A:Cross-references: EMBL:AF157706; PIDN:AAD49652.1
 A:Experimental source: strain 229; variant B
 C:Genetics:
 A:Note: U38
 C:Superfamily: herpesvirus DNA-directed DNA polymerase
 C:Keywords: nucleotidyltransferase

Query Match 1.0%; Score 8; DB 2; Length 1012;
 Best Local Similarity 100.0%; Pred. No. 50;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 574 VIYGDPTS 581
 |||||
 DB 736 VIYGDPTS 743

RESULT 56
 J01970
 DNA-directed DNA polymerase (EC 2.7.7.7) - Lymantria dispar nuclear polyhedrosis virus
 C:Species: Lymantria dispar nuclear polyhedrosis virus, LdNPV
 C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 20-Jun-2000
 C:Accession: J01970
 R:Johnson, R.M.; Glocker, B.; Rohmann, G.F.
 J: Gen. Virol. 73, 3177-3183, 1992
 A:Title: Characterization of the nucleotide sequence of the Lymantria dispar nuclear polyhedrosis virus
 A:Reference number: P00515; MUID:93107852; PMID:1469355
 A:Accession: J01970
 A:Molecule type: DNA
 A:Residues: 1-1013 <BD>
 A:Cross-references: DDBJ:D11476; NID:G1041123; PIDN:BA02036.1; PID:G222230
 C:Superfamily: herpesvirus DNA-directed DNA polymerase
 C:Keywords: DNA binding; nucleotidyltransferase

Query Match 1.0%; Score 8; DB 2; Length 1013;
 Best Local Similarity 100.0%; Pred. No. 50;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 574 VIYGDPTS 581
 |||||
 DB 682 VIYGDPTS 689

RESULT 57
 T30431
 DNA-directed DNA polymerase (EC 2.7.7.7) - Lymantria dispar nuclear polyhedrosis virus
 C:Species: Lymantria dispar nuclear polyhedrosis virus, LdNPV
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jul-2000
 C:Accession: T30431
 R:Kuzio, V.; Pearson, M.N.; Harwood, S.H.; Funk, C.J.; Evans, J.T.; Slavicek, J.M.; Rohr
 Virol. 253, 17-34, 1999
 A:Title: Sequence and analysis of the genome of a baculovirus pathogenic for Lymantria d
 A:Reference number: Z20836; MUID:99124785; PMID:9887215
 A:Accession: T30431
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1014 <KUZ>
 A:Cross-references: EMBL:AF061810; NID:G3822234; PIDN:AAC70269.1; PID:G3822318
 C:Superfamily: herpesvirus DNA-directed DNA polymerase
 C:Keywords: DNA binding; nucleotidyltransferase

Query Match 1.0%; Score 8; DB 2; Length 1014;
 Best Local Similarity 100.0%; Pred. No. 50;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 574 VIYGDPTS 581
 |||||
 DB 682 VIYGDPTS 689

RESULT 58
 T03108
 DNA-directed DNA polymerase (EC 2.7.7.7) - alcelaphine herpesvirus 1
 C:Species: alcelaphine herpesvirus 1
 C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 18-Jun-1999
 C:Accession: T03108
 R:Enser, A.; Pelanz, R.; Fleckenstein, B.
 J: Virol. 71, 6517-6525, 1997
 A:Title: Primary structure of the alcelaphine herpesvirus 1 genome.
 A:Reference number: Z14840; MUID:97404659; PMID:9261371
 A:Accession: T03108
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1026 <ENS>
 A:Cross-references: EMBL:AF005370; NID:G2337967; PIDN:AC58060.1; PID:G2337976
 C:Superfamily: herpesvirus DNA-directed DNA polymerase
 C:Keywords: DNA binding; DNA replication; nucleotidyltransferase

Query Match 1.0%; Score 8; DB 2; Length 1026;
 Best Local Similarity 100.0%; Pred. No. 50;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 576 YGDTSLF 583
 |||||
 DB 770 YGDTSLF 777

RESULT 59
 S19661
 DNA-directed DNA polymerase (EC 2.7.7.7) III large chain - fission yeast (Schizosaccharo
 C:Species: Schizosaccharomyces pombe
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: S19661
 R:Pignede, G.; Bouvier, D.; de Recondo, A.M.; Baladaci, G.
 J: Mol. Biol. 222, 209-218, 1991
 A:Title: Characterization of the POB3 gene product from Schizosaccharomyces pombe indica
 A:Reference number: S19661; MUID:92071854; PMID:11960723
 A:Accession: S19661
 A:Molecule type: DNA
 A:Residues: 1-1084 <PIG>
 A:Cross-references: EMBL:X59278; NID:G5010; PIDN:CAA4166.1; PID:G5011
 C:Genetics: 77/1
 C:Superfamily: herpesvirus DNA-directed DNA polymerase
 C:Keywords: DNA binding; nucleotidyltransferase; nucleus

Query Match 1.0%; Score 8; DB 1; Length 1084;
 Best Local Similarity 100.0%; Pred. No. 53;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 574 VIYGDPTS 581
 |||||
 DB 737 VIYGDPTS 744

RESULT 60
 T43266
 DNA-directed DNA polymerase (EC 2.7.7.7) delta chain - fission yeast (Schizosaccharomyce
 N:Alternate names: DNA polymerase delta
 C:Species: Schizosaccharomyces pombe
 C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 21-Jul-2000
 C:Accession: T43266
 R:Pignede, G.; Bouvier, D.; de Recondo, A.M.; Baladaci, G.
 J: Mol. Biol. 222, 209-218, 1991

A>Title: Characterization of the POL3 gene product from Schizosaccharomyces pombe indic
 A/Reference number: S19661; MUID:92071954; PMID:1960723
 A/Accession: T43266
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-1086 <PIG>
 A/Cross-references: EMBL:L07734; NID:g173383; PIDN:AAA35303.1; PID:g173384
 C/Genetics:
 A/Gene: pol3
 A/Introns: 77/1
 C/Superfamily: herpesvirus DNA-directed DNA polymerase
 C/Keywords: DNA binding; DNA replication; nucleotidyltransferase

Query Match 1.0%; Score 8; DB 2; Length 1086;
 Best Local Similarity 100.0%; Pred. No. 53;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 574 VIYGDTS 581
 |||||
 737 VIYGDTS 744

RESULT 61
 T40242
 DNA polymerase delta large chain - fission yeast (Schizosaccharomyces pombe)
 C/Species: Schizosaccharomyces pombe
 C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
 C/Accession: T40242
 R/Bozrym, K.; Beck, A.; Reinhardt, R.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.
 submitted to the EMBL Data Library, October 1999
 A/Reference number: Z21916
 A/Accession: T40242
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-1086 <BOR>
 A/Cross-references: EMBL:AL121815; PIDN:CAB58156.1; GSPDB:GN00667; SPDB:SPBC336.04
 A/Experimental source: strain 972h-; cosmid c336
 C/Genetics:
 A/Gene: SPDB:SPBC336.04
 A/Map position: 2
 A/Introns: 77/1
 C/Superfamily: herpesvirus DNA-directed DNA polymerase

Query Match 1.0%; Score 8; DB 2; Length 1086;
 Best Local Similarity 100.0%; Pred. No. 53;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

574 VIYGDTS 581
 |||||
 Db 737 VIYGDTS 744

RESULT 62
 T05731
 DNA-directed DNA polymerase (EC 2.7.7.7) delta chain - soybean
 C/Species: Glycine max (soybean)
 C/Date: 09-Apr-1999 #sequence_revision 09-Apr-1999 #text_change 18-Jun-1999
 C/Accession: T05731
 R/Collins, J.T.B.; Cannon, G.C.; Heinrich, S.
 submitted to the EMBL Data Library, August 1997
 A/Reference number: Z15439
 A/Accession: T05731
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-1088 <COL>
 A/Cross-references: EMBL:AF020193; NID:g2895197; PIDN:AAC18443.1; PID:g2895198
 C/Genetics:
 A/Gene: Pol delta
 C/Function:
 A/Description: catalyzes replication of DNA
 C/Superfamily: herpesvirus DNA-directed DNA polymerase
 C/Keywords: DNA binding; DNA replication; nucleotidyltransferase

Query Match 1.0%; Score 8; DB 2; Length 1088;
 Best Local Similarity 100.0%; Pred. No. 53;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 574 VIYGDTS 581
 |||||
 Db 714 VIYGDTS 721

RESULT 63
 S22573
 DNA-directed DNA polymerase (EC 2.7.7.7) delta - malaria parasite (Plasmodium falciparum)
 C/Species: Plasmodium falciparum
 C/Date: 22-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 09-Jun-2000
 C/Accession: S22573; A45639; S17330
 R/Ridley, R.G.; White, J.H.; McAlleese, S.M.; Goman, M.; Alano, P.; de Vries, E.; Kilbey,
 Nucleic Acids Res. 19, 6731-6736, 1991
 A/Title: DNA polymerase delta: gene sequences from Plasmodium falciparum indicate that
 A/Reference number: S22573; MUID:92107655; PMID:1762904
 A/Accession: S22573
 A/Molecule type: DNA
 A/Residues: 1-1094 <RID>
 A/Cross-references: EMBL:X62423; NID:g99940; PIDN:CAA44289.1; PID:g9941
 R/Fox, B.A.; Bzik, D.J.
 Mol. Biochem. Parasitol. 49, 289-296, 1991

A/Title: The primary structure of Plasmodium falciparum DNA polymerase delta is similar
 A/Reference number: A45639; MUID:92131065; PMID:1775172
 A/Accession: A45639
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-1094 <FOX>
 A/Cross-references: GB:M64715; NID:g160269; PIDN:AAA29589.1; PID:g160270
 A/Note: sequence extracted from NCBI backbone (NCBIN:77767, NCBIPI:77769)
 C/Superfamily: herpesvirus DNA-directed DNA polymerase
 C/Keywords: DNA binding; DNA replication; exonuclease; nucleotidyltransferase; nucleus;
 F:1003-1019/Region: zinc finger CCCC motif
 F:1049-1067/Region: zinc finger CCCC motif

Query Match 1.0%; Score 8; DB 2; Length 1094;
 Best Local Similarity 100.0%; Pred. No. 53;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 574 VIYGDTS 581
 |||||
 Db 749 VIYGDTS 756

RESULT 64
 S40243
 DNA-directed DNA polymerase (EC 2.7.7.7) delta chain - mouse
 C/Species: Mus musculus (house mouse)
 C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C/Accession: S40243; J70670
 R/Cullmann, G.; Hindges, R.; Bertschold, M.W.; Huebscher, U.
 submitted to the EMBL Data Library, March 1993
 A/Reference number: S40243
 A/Accession: S40243
 A/Molecule type: mRNA
 A/Residues: 1-1105 <CUH>
 A/Cross-references: EMBL:Z21848; NID:g438133; PIDN:CAA78895.1; PID:g438134
 R/Cullmann, G.; Hindges, R.; Bertschold, M.W.; Huebscher, U.
 Gene 134, 191-200, 1993
 A/Title: Cloning of a mouse cDNA encoding DNA polymerase delta: refinement of the homolo
 A/Reference number: J70670; MUID:94085777; PMID:8262377
 A/Accession: J70670
 A/Molecule type: DNA
 A/Residues: 1-1111, 'G', '113', 'P', '115-1034', 'Y', '1036-1105' <CU2>
 A/Cross-references: EMBL:Z21848
 A/Note: the sequence translated from Z21848 is inconsistent with that from this sequence
 C/Comment: Three DNA polymerases alpha, delta and epsilon chains are essential. This de
 C/Genetics:
 A/Gene: poldelta
 C/Superfamily: herpesvirus DNA-directed DNA polymerase

C/Keywords: DNA replication; heterodimer; nucleotidyltransferase

Query Match 1.0%; Score 8; DB 1; Length 1105;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 621 KRRYAGLL 628
|||
Db 804 KRRYAGLL 811

RESULT 65

A39299
DNA-directed DNA polymerase (EC 2.7.7.7) delta catalytic chain - bovine

N/Alternate names: DNA-directed DNA polymerase III 125K chain

N/Comments: DNA-directed DNA polymerase delta, 116K chain

C/Species: Bos primigenius taurus (cattle)

C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C/Accession: A39299

R/Chang, J.; Chung, D.W.; Tan, C.K.; Downey, K.M.; Davie, E.W.; So, A.G.

R/Chemistry 30, 11742-11750, 1991

A/Title: Primary structure of the catalytic subunit of calf thymus DNA polymerase delta;

A/Reference number: A39299; PMID:92089082; PMID:1721537

A/Accession: A39299

A/Molecule type: mRNA

A/Residues: 1-1106 <ZHA>

A/Cross-references: GB:M80395; NID:G162973; PIDN:AAA30493.1; PID:G162974

A/Experimental source: thymus

A/Note: 131-Glu and 263-Ser were also found by mRNA sequencing

A/Note: part of this sequence was confirmed by protein sequencing

A/Note: The amino end of the mature protein was blocked

C/Comment: A 116K polypeptide was shown to be a degradation product of the mature protein

C/Superfamily: herpesvirus DNA-directed DNA polymerase

C/Keywords: DNA binding; DNA replication; nucleotidyltransferase; nucleus; zinc finger

Query Match 1.0%; Score 8; DB 1; Length 1106;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 621 KRRYAGLL 628
|||
Db 805 KRRYAGLL 812

RESULT 66

A41618

DNA-directed DNA polymerase (EC 2.7.7.7) delta catalytic chain - human

C/Species: Homo sapiens (man)

C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C/Accession: A41618; S35455

R/Chung, D.W.; Zhang, J.; Tan, C.K.; Davie, E.W.; So, A.G.; Downey, K.M.

R/Proc. Natl. Acad. Sci. U.S.A. 88, 11197-11201, 1991

A/Title: Primary structure of the catalytic subunit of human DNA polymerase delta and cH

A/Reference number: A41618; PMID:92107916; PMID:1722322

A/Accession: A41618

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-1107 <CHU>

A/Cross-references: GB:M80397; NID:G181619; PIDN:AAA58439.1; PID:G181620

R/Yang, C.L.; Chang, L.S.; Zhang, P.; Hao, H.; Zhu, L.; Toomey, N.L.; Lee, M.Y.W.T.

Nucleic Acids Res. 20, 735-745, 1992

A/Title: Molecular cloning of the cDNA for the catalytic subunit of human DNA polymerase

A/Reference number: S35455; PMID:92178967; PMID:1542570

A/Accession: S35455

A/Molecule type: mRNA

A/Residues: 1-29, 'R', 31-118, 'H', 120-172, 'N', 174-471, 'V', 473-775, 'G', 777-1107 <YAN>

A/Cross-references: EMBL:M81755; NID:G181621; PIDN:AAA35768.1; PID:G181622

C/Genetics

A/Genes: GDB:POLDI, POLD

A/Cross-references: GDB:129089; OMIM:174761

A/Map Position: 19q13.3-19q13.3

C/Superfamily: herpesvirus DNA-directed DNA polymerase

C/Keywords: DNA binding; DNA replication; nucleotidyltransferase; nucleus; zinc finger

Query Match 1.0%; Score 8; DB 1; Length 1107;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 621 KRRYAGLL 628
|||
Db 806 KRRYAGLL 813

RESULT 67

DJ8633

DNA-directed DNA polymerase (EC 2.7.7.7) - equine herpesvirus 1 (strain Ab4p)

C/Species: equine herpesvirus 1

A/Note: host Equus caballus (domestic horse)

C/Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 11-Jun-1999

C/Accession: D36798

R/Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.

submitted to GenBank, March 1992

A/Description: The DNA sequence of equine herpesvirus-1.

A/Reference number: A36805

A/Accession: D36798

A/Molecule type: DNA

A/Residues: 1-1220 <TEL>

A/Cross-references: GB:M86664; NID:G330791; PIDN:AA02465.1; PID:G330822

R/Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.

Virology 189, 304-316, 1992

A/Title: The DNA sequence of equine herpesvirus-1.

A/Reference number: A41831; PMID:9229566; PMID:1318606

A/Contents: annotation; possible protein-coding frames

A/Note: neither amino acid nor nucleotide sequence is given

C/Genetics

A/Genes: 30

C/Superfamily: herpesvirus DNA-directed DNA polymerase

C/Keywords: DNA binding; DNA replication; nucleotidyltransferase

Query Match 1.0%; Score 8; DB 1; Length 1220;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 574 VVIGDTRDS 581
|||
Db* 883 VVIGDTRDS 890

RESULT 68

T42573

DNA-directed DNA polymerase (EC 2.7.7.7) - equine herpesvirus 4 (strain NS80567)

C/Species: equine herpesvirus 4

A/Variety: strain NS80567

C/Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 21-Jul-2000

C/Accession: T42573

R/Telford, E.A.R.; Watson, M.S.; Perry, J.; Cullinane, A.A.; Davison, A.J.

J. Gen. Virol. 79, 1197-1203, 1998

A/Title: The DNA sequence of equine herpesvirus-4.

A/Reference number: 222173; PMID:98264497; PMID:9603335

A/Accession: T42573

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-1220 <TEL>

A/Cross-references: EMBL:AF030027; NID:G2605950; PIDN:AA059546.1; PID:G2605974

A/Experimental source: strain NS80567

C/Genetics

A/Genes: 30

C/Superfamily: herpesvirus DNA-directed DNA polymerase

C/Keywords: DNA binding; DNA replication; nucleotidyltransferase

Query Match 1.0%; Score 8; DB 2; Length 1220;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 574 VVIGDTRDS 581
|||

Db 884 VIYGDTS 891

RESULT 69

DUBEC1

DNA-directed DNA polymerase (EC 2.7.7.7) - human cytomegalovirus (strain AD169)
N:Alternate names: HPLF2 protein; UL54 protein
C:Species: human cytomegalovirus, human herpesvirus 5
A:Note: host Homo sapiens (man)C>Date: 31-Mar-1988 #sequence_revision 31-Dec-1990 #text_change 11-Jun-1999
C/Accession: S09817; A25983

R:Chee, M.S.; Bankier, A.T.; Beck, S.; Bohni, R.; Brown, C.M.; Cerny, R.; Hornell, T.; M.; Barrell, B.G.

Curr. Top. Microbiol. Immunol. 154, 125-169, 1990

A>Title: Analysis of the protein-coding content of the sequence of human cytomegalovirus
A/Reference number: S09749; MUID:90269039; PMID:2161319

A/Accession: S09817

A/Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-1242 <CHE>

A/Cross-references: EMBL:X17403; NID:G59591; PIDN:CAA5413.1; PID:G1780832

A/Note: possible protein-coding frames are given

A/Note: the DNA sequence was submitted to EMBL, December 1989, in computer-readable form

R:Kouzarides, T.; Bankier, A.T.; Satchwell, S.C.; Weston, K.; Tomlinson, P.; Barrell, B.

J. Virol. 61, 125-133, 1987

A>Title: Sequence and transcription analysis of the human cytomegalovirus DNA polymerase

A/Reference number: A25983; MUID:87061230; PMID:3023650

A/Accession: A25983

A/Molecule type: DNA

A/Residues: 1-1242 <KOU>

A/Cross-references: GB:M14709; NID:G330640; PIDN:AAA45988.1; PID:G330642

C/Genetics:

A/Map position: unique long region (U-L)

C/Superfamily: herpesvirus DNA-directed DNA polymerase

C/Keywords: DNA binding; DNA replication; nucleotidyltransferase

Query Match

Best Local Similarity 1.0%; Score 8; DB 1; Length 1242;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 574 VIYGDTS 581

Db 906 VIYGDTS 913

RESULT 70

A: fiber protein gp37 - phage K3

C/Species: phage K3

C/Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 20-Sep-1999

C/Accession: S07278; I41090; I41091

R:Riede, I.; Drexler, K.; Eschbach, M.L.; Henning, U.

J. Mol. Biol. 191, 255-266, 1996

A>Title: DNA sequence of the tail fiber genes 37, encoding the receptor recognizing part

A/Reference number: S07278; MUID:87112716; PMID:3806672

A/Accession: S07278

A/Molecule type: DNA

A/Residues: 1-1243 <RIE>

A/Cross-references: EMBL:X04747; NID:G15110; PIDN:CAA28445.1; PID:G15111

R:Riede, I.; Eschbach, M.L.; Henning, U.

Mol. Gen. Genet. 195, 144-152, 1984

A>Title: DNA sequence heterogeneity in the genes of T-even type Escherichia coli phages

A/Reference number: I41090; MUID:85035820; PMID:6052843

A/Accession: I41090

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 158-211 <RES>

A/Cross-references: EMBL:X00613; NID:G41844; PIDN:CAA25250.1; PID:G929577

A/Accession: I41091

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 727-792 <RE2>

A/Cross-references: EMBL:X00615; NID:G41845; PIDN:CAA25251.1; PID:G929578

C/Genetics:
A/Genes: 37
C/Superfamily: phage T4 tail fiber protein gp37
C/Keywords: dimer; tail fiber

Query Match 1.0%; Score 8; DB 2; Length 1243;

Best Local Similarity 100.0%; Pred. No. 60;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 VLPEAVRE 102

Db 1196 VLPEAVRE 1203

RESULT 71

S20052

DNA-directed DNA polymerase (EC 2.7.7.7) alpha catalytic chain - Trypanosoma brucei brucei

C/Species: Trypanosoma brucei brucei

C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C/Accession: S20052; S15729

R:Deegater, P.A.J.; Strating, M.; Murphy, N.B.; Kooy, R.F.; van der Vliet, P.C.; Overdu

Nucleic Acids Res. 19, 6441-6447, 1991

A>Title: The Trypanosoma brucei DNA polymerase alpha core subunit gene is developmentally

A/Reference number: S20051; MUID:92093600; PMID:1754381

A/Accession: S20052

A/Molecule type: DNA

A/Residues: 1-1339 <LEB>

A/Cross-references: EMBL:X60951; NID:G10499; PIDN:CAA43287.1; PID:G10501

A/Note: 828-Arg was also found

C/Genetics:

A/Genes: pola

C/Superfamily: herpesvirus DNA-directed DNA polymerase

C/Keywords: DNA binding; DNA replication; nucleotidyltransferase; nucleus; zinc finger

Query Match 1.0%; Score 8; DB 1; Length 1339;

Best Local Similarity 100.0%; Pred. No. 64;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 574 VIYGDTS 581

Db 894 VIYGDTS 901

RESULT 72

JC5508

DNA-directed DNA polymerase (EC 2.7.7.7) alpha catalytic chain - Leishmania donovani

C/Species: Leishmania donovani

C/Date: 02-Sep-1997 #sequence_revision 05-Sep-1997 #text_change 18-Jun-1999

C/Accession: JC5508

R:Luton, K.; Johnson, A.M.

Biochem. Biophys. Res. Commun. 234, 95-100, 1997

A>Title: Cloning and sequence analysis of the DNA polymerase alpha gene of Leishmania dc

A/Reference number: JC5508; MUID:97312527; PMID:9168968

A/Accession: JC5508

A/Molecule type: DNA

A/Residues: 1-1339 <LUT>

A/Cross-references: GB:U78172; NID:G2138287; PIDN:AAQ47538.1; PID:G2138288

C/Superfamily: herpesvirus DNA-directed DNA polymerase

C/Keywords: nucleotidyltransferase

Query Match 1.0%; Score 8; DB 2; Length 1339;

Best Local Similarity 100.0%; Pred. No. 64;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 574 VIYGDTS 581

Db 896 VIYGDTS 903

RESULT 73

S09579

tail fiber protein gp37 - phage T2

C/Species: phage T2

C/Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 20-Sep-1999
 C/Accession: S09579
 R/Release: I.; Drexler, K.; Eschbach, M.L.; Henning, U.
 J. Mol. Biol. 191, 255-266, 1986
 A/Title: DNA sequence of the tail fiber genes 37, encoding the receptor recognizing part
 A/Reference number: S07278; MUID:87112716; PMID:3806672
 A/Accession: S09579
 A/Molecule type: DNA
 A/Residues: 11341 <RES>
 A/Cross-references: EMBL:X04442; NID:915195; PIDN:CAA28038.1; PID:915196
 C/Genetics:
 A/Genes: 37
 C/Superfamily: phage T4 tail fiber protein gp37
 C/Keywords: dimer; tail fiber

Query Match 1.0%; Score 8; DB 2; Length 1341;
 Best Local Similarity 100.0%; Pred. No. 64;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 VLPEAVRE 102
 |||||
 1294 VLPEAVRE 1301

RESULT 74
 DJZPA
 DNA-directed DNA polymerase (EC 2.7.7.7) alpha - fusion yeast (Schizosaccharomyces pombe
 C/Species: Schizosaccharomyces pombe
 C/Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 10-Dec-1999
 C/Accession: S15993; F38752; S21818
 R/Dumaner, V.; Tiller, J.; de Rendo, A.M.; Baldacci, G.
 Mol. Gen. Genet. 226, 182-189, 1991
 A/Title: The POL1 gene from the fusion yeast, Schizosaccharomyces pombe, shows conserve
 A/Reference number: S15993; MUID:91238692; PMID:2034212
 A/Accession: S15993
 A/Molecule type: DNA
 A/Residues: 1-1405 <MOI>
 A/Cross-references: EMBL:X58299; NID:95008; PIDN:CAA1232.1; PID:95009
 R/Oliver, K.; Harris, D.; Wood, V.; Bartell, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, September 1996
 A/Reference number: 221809
 A/Accession: T38752
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-1405 <OLI>
 A/Cross-references: EMBL:Z99296; PIDN:CAB16598.1; GSPDB:GNO0066; SPDB:SPAC3H5.06C
 A/Experimental source: strain 972h-; comid c3H5
 C/Genetics: <MOI>
 Gene: POL1
 Map position: II
 Genetics: <OLI>
 A/Genes: SPAC3H5.06C
 A/Map position: 1
 A/Intons: 101/1
 C/Superfamily: herpesvirus DNA-directed DNA polymerase
 C/Keywords: DNA binding; DNA replication; nucleotidyltransferase; zinc finger
 F,116-1340/Region: zinc finger CCCC motif

Query Match 1.0%; Score 8; DB 1; Length 1405;
 Best Local Similarity 100.0%; Pred. No. 66;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 574 VIYGDYDS 581
 |||||
 Db 978 VIYGDYDS 985

RESULT 75
 DJHUAC
 DNA-directed DNA polymerase (EC 2.7.7.7) alpha catalytic chain - human
 C/Species: Homo sapiens (man)
 C/Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 11-Jun-1999
 C/Accession: S00257; A30440; T57513; S12665

R/Mong, S.W.; Wahl, A.F.; Yuan, P.M.; Arai, N.; Pearson, B.E.; Arai, K.I.; Korn, D.; Hu
 EXMO 7, 7, 37-47, 1988
 A/Title: Human DNA polymerase alpha gene expression is cell proliferation dependent and
 A/Reference number: S00257; MUID:88196090; PMID:3339994
 A/Accession: S00257
 A/Molecule type: mRNA
 A/Residues: 1-1462 <MON>
 A/Cross-references: EMBL:X06745; NID:935567; PIDN:CAA29920.1; PID:935568
 A/Accession: A30440
 A/Molecule type: protein
 A/Residues: 438-449; 495-502; 'G', 838-848; 1090-1105; 1201-1216; 1397-1407; 1444-1453 <MON2>
 R/Pearson, B.E.; Nashner, H.P.; Wang, T.S.
 Mol. Cell. Biol. 11, 2081-2095, 1991
 A/Title: Human DNA polymerase alpha gene: sequences controlling expression in cycling an
 A/Reference number: 157513; MUID:91172197; PMID:2005899
 A/Accession: 157513
 A/Status: translation not shown; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-8 <RES>
 A/Cross-references: GB:M64481; NID:9181617; PIDN:AA52318.1; PID:9181618
 R/Hel, K.L.; Copeland, W.C.; Wang, T.S.F.
 Nucleic Acids Res. 18, 6231-6237, 1990
 A/Title: Human DNA polymerase alpha catalytic polypeptide binds ConA and contain
 A/Reference number: S12665; MUID:91057099; PMID:2243771
 A/Accession: S12665
 A/Molecule type: protein
 A/Residues: 19-37; 'C', 1406-1425, 'C' <HSI>
 C/Genetics:
 A/Genes: GDB:POLA
 A/Cross-references: GDB:120304; OMIM:312040
 A/Map position: XP22.3-Xp21.1
 C/Superfamily: herpesvirus DNA-directed DNA polymerase
 C/Keywords: DNA binding; DNA replication; nucleotidyltransferase; zinc finger
 F,1249-1374/Region: zinc fingers

Query Match 1.0%; Score 8; DB 1; Length 1462;
 Best Local Similarity 100.0%; Pred. No. 69;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 574 VIYGDYDS 581
 |||||
 Db 998 VIYGDYDS 1005

Search completed: April 23, 2003, 07:50:01
 Job time: 48 secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 07:28:12 ; Search time 29 Seconds

(without alignments)
814.711 Million cell updates/sec

Title: US-10-034-849-2

Perfect score: 4188

Sequence: 1 MTEVFTVLDSSEVVGKEP.....LKASATGQKTLFPFLAKKSK 803

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents AA:
1: /cgn2_6/ptodata/1/1aa/5A COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5a COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/6C COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4188	100.0	803	2	US-08-907-166-4
2	3028.5	72.3	803	1	US-08-062-368-4
3	2963.5	70.8	803	1	US-08-062-368-2
4	1563	37.3	788	2	US-08-907-166-6
5	1237.5	29.5	774	2	US-08-902-632-2
6	1237.5	29.5	776	2	US-08-688-649-37
7	1236.5	29.5	775	1	US-07-966-278-1
8	1236.5	29.5	775	1	US-08-424-921-1
9	1236.5	29.5	775	2	US-08-556-355A-1
10	1236.5	29.5	775	2	US-07-803-627A-1
11	1234.5	29.5	774	3	US-03-073-354-1
12	1234.5	29.5	774	3	US-08-656-005A-1
13	1234.5	29.5	774	4	US-09-073-259-1
14	1234.5	29.5	774	4	US-09-363-095-1
15	1234.5	29.5	774	4	US-09-418-027-1
16	1146.5	27.4	778	2	US-08-906-925-4
17	1104	26.4	779	1	US-08-375-134-12
18	1104	26.4	779	5	PCT-US95-15263-12
19	760.5	18.2	1107	5	US-08-366-577-2
20	760.5	18.2	1107	5	PCT-US96-00005-2
21	739.5	17.7	1019	1	US-08-271-364A-7
22	739.5	17.7	1019	2	US-08-223-715B-26
23	635	15.2	1022	1	US-08-271-364A-8
24	635	15.2	1022	2	US-08-223-715B-27
25	579.5	13.8	1462	3	US-07-792-600-31
26	579.5	13.8	1462	3	US-09-157-021-31
27	579.5	13.8	1462	3	US-09-156-842-31

28	542.5	13.0	877	2	US-08-907-166-8	Sequence 8, Appl
29	524.5	12.5	1015	2	US-08-680-326-32	Sequence 32, Appl
30	515	12.3	1009	2	US-08-680-326-31	Sequence 31, Appl
31	506.5	12.1	762	2	US-08-907-166-10	Sequence 10, Appl
32	502.5	12.0	837	2	US-08-680-326-117	Sequence 117, App
33	495	11.8	1008	2	US-08-680-326-30	Sequence 30, Appl
34	458.5	10.9	396	1	US-08-229-284A-2	Sequence 2, Appl
35	452	10.8	1240	2	US-08-680-326-37	Sequence 37, Appl
36	450.5	10.8	1220	2	US-08-680-326-38	Sequence 38, Appl
37	450	10.7	1194	2	US-08-680-326-35	Sequence 35, Appl
38	444	10.6	1235	2	US-08-680-326-36	Sequence 36, Appl
39	440.5	10.5	783	1	US-08-101-593-6	Sequence 6, Appl
40	440.5	10.5	783	2	US-08-465-995A-6	Sequence 6, Appl
41	440.5	10.5	783	2	US-08-465-994C-6	Sequence 6, Appl
42	424	10.1	1012	2	US-08-680-326-34	Sequence 34, Appl
43	419	10.0	1094	2	US-08-680-326-40	Sequence 40, Appl
44	394	9.4	1097	2	US-08-680-326-39	Sequence 39, Appl
45	368	8.8	1242	2	US-08-680-326-33	Sequence 33, Appl

ALIGNMENTS

RESULT 1									
US-08-907-166-4									
; Sequence 4, Application US/08907166									
; Patent No. 5948666									
; GENERAL INFORMATION:									
; APPLICANT: Callen, Walter									
; TITLE OF INVENTION: ISOLATION AND IDENTIFICATION OF NOVEL POLYMERASES									
; FILE REFERENCE: 09010/027001									
; CURRENT APPLICATION NUMBER: US/08/907,166									
; CURRENT FILING DATE: 1997-08-06									
; NUMBER OF SEQ ID NOS: 12									
; SOFTWARE: FastSeq for Windows Version 3.0									
; SEQ ID NO 4									
; LENGTH: 803									
; TYPE: PRT									
; ORGANISM: Pyrobolus fumarius									
US-08-907-166-4									
Query Match									
Best Local Similarity 100.0%; Score 4188; DB 2; Length 803;									
Matches 803; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	MTEVFTVLDSSEVVGKEPQVILMGIAENGSRVLLDRSFRFYFALLAPGADPKQVAQ	60						
Db	1	MTEVFTVLDSSEVVGKEPQVILMGIAENGSRVLLDRSFRFYFALLAPGADPKQVAQ	60						
Qy	61	RIRALSPKSPITIGVEDDKRYFGRRVLRIRTVLPEAVREYRELVKNVGVEDVLEAD	120						
Db	61	RIRALSPKSPITIGVEDDKRYFGRRVLRIRTVLPEAVREYRELVKNVGVEDVLEAD	120						
Qy	121	IRFMRRLIDHDLPPFWYVVEAPLENKMGFRVDKRYLVKSRPEPIYGGALPTKLPDL	180						
Db	121	IRFMRRLIDHDLPPFWYVVEAPLENKMGFRVDKRYLVKSRPEPIYGGALPTKLPDL	180						
Qy	181	RILAFDLEVYSKQSPRPEDRPVILAVKTDGDEVFLIAGKODRKPIEFPEYVGRYD	240						
Db	181	RILAFDLEVYSKQSPRPEDRPVILAVKTDGDEVFLIAGKODRKPIEFPEYVGRYD	240						
Qy	241	PDIIVGNNHFPDMPYLRLRRARILIGIKLQDVTRVGAEPITTSVGHVSPGRLNVDLYDYA	300						
Db	241	PDIIVGNNHFPDMPYLRLRRARILIGIKLQDVTRVGAEPITTSVGHVSPGRLNVDLYDYA	300						
Qy	301	EEWEIKIKLEAEVLAELVGMKKSERYITIMWEIIPVDWDPKRRPLILQYARDVATYGS	360						
Db	301	EEWEIKIKLEAEVLAELVGMKKSERYITIMWEIIPVDWDPKRRPLILQYARDVATYGS	360						
Qy	361	LAETILPFAQLSVTGLPLPDQVAMSVGRLEEMYLIRAFKPKKEVLPNVEPEETRYG	420						
Db	361	LAETILPFAQLSVTGLPLPDQVAMSVGRLEEMYLIRAFKPKKEVLPNVEPEETRYG	420						

Qy 421 AIVLEPLRGVHENIAVLDPSSMYPMIKYVNPDTLVPRGKCGCGCWAPEVKHFR 480
Db 421 AIVLEPLRGVHENIAVLDPSSMYPMIKYVNPDTLVPRGKCGCGCWAPEVKHFR 480
Qy 481 RCPGPKVTVLELLELRKRVAAEMKTPPSPPEYRLDROQALKVLANASYGMGSG 540
Db 481 RCPGPKVTVLELLELRKRVAAEMKTPPSPPEYRLDROQALKVLANASYGMGSG 540
Qy 541 ARMYCBKAAVTAMGRHLIRTAIIRAKGLKVIYDGTDSLFTVYDPEKVENFKIIE 600
Db 541 ARMYCBKAAVTAMGRHLIRTAIIRAKGLKVIYDGTDSLFTVYDPEKVENFKIIE 600
Qy 601 ELGPEIKLEKVKYRLLFTEAKKRYAGLLEDRIDIVGFEAVRGDMCELAKEVQTVKEIV 660
Db 601 ELGPEIKLEKVKYRLLFTEAKKRYAGLLEDRIDIVGFEAVRGDMCELAKEVQTVKEIV 660
Qy 661 LKTSVNVKAVEYRKIVKELEEGKVPKEKVIYMKTLISKRLSEYTTAPHVAAKRMLSAG 720
Db 661 LKTSVNVKAVEYRKIVKELEEGKVPKEKVIYMKTLISKRLSEYTTAPHVAAKRMLSAG 720
Qy 721 YRVSPPGKIGYIVVKGGRISORAMPYPMVNDPSOLIDVTYVDHOIIPALRIIGYGIT 780
Db 721 YRVSPPGKIGYIVVKGGRISORAMPYPMVNDPSOLIDVTYVDHOIIPALRIIGYGIT 780
Qy 781 EKLKASATGOKTLPDLFLAKSK 803
Db 781 EKLKASATGOKTLPDLFLAKSK 803

RESULT 2

US-08-062-368-4
Sequence 4, Application US/08062368
Patent No. 5491086
GENERAL INFORMATION:
APPLICANT: Gelfand, David H.
TITLE OF INVENTION: Purified Thermostable Nucleic Acid
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Hoffmann-La Roche Inc.
STREET: 340 Kingsland Street
CITY: Nutley
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 07110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/062,368
FILING DATE: 19930514
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sias Ph.D., Stacey R.
REGISTRATION NUMBER: 32,630
REFERENCE/DOCKET NUMBER: 8584
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 814-2863
TELEFAX: (510) 814-2877
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 803 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-062-368-4

Query Match

72.3%, Score 3028.5, DB 1, Length 803;

Best Local Similarity 70.8%; Pred. No. 1e-270;
Matches 570; Conservative 106; Mismatches 122; Indels 7; Gaps 5;

Qy 1 MTEVV-FTVLDSSYEVVGEKPOVITWGLAENGERRVTLIDNSFRFPYALALAGAD--KQ 57
Db 1 MTEVEFLVLDSSYEVVGEKPOVITWGLAENGERRVTLIDNSFRFPYALALAGAD--KQ 57
Qy 58 VAQRIRALSRRKSPDIIGVEDDKRYGPRRLVIRTVLPEAAVEAEVLEKVNVDGEDVL 117
Db 61 IASIRRLSVKSPDIIGVEDDKRYGPRRLVIRTVLPEAAVEAEVLEKVNVDGEDVL 120
Qy 118 EADIRFARVYLDIDLPFTFTVREAEPELENMGRVNDKVLVYSREPEYALAPTL 177
Db 121 EADIRFARVYLDIDLPFTFTVREAEPELENMGRVNDKVLVYSREPEYALAPTL 178
Qy 178 PDRLIAPDIDVYSKQSPERDPVTVIAVKTDDGDEVLFAEGKDRKPIREFEYVK 237
Db 179 PEMRLVADIDVYSRGRSPNADPVIIVSLDSSGKERLLEAGHDRAVLAEFEYVR 238
Qy 238 RYDPDIIYGNMNPMPYLLRPAARILGKLDYPRVGAEPETTSYGHVSVGRVLDVLY 297
Db 239 ADPDIIVGNMNPMPYLLRPAARILGKLDYPRVGAEPETTSYGHVSVGRVLDVLY 298
Qy 298 DVAEEMPEIKLSLEVAEVLGYMKSGSERVITINMEIDYVDDPKRPLLYARDVRA 357
Db 299 DVAEEMPEIKLSLEVAEVLGYMKSGSERVITINMEIDYVDDPKRPLLYARDVRA 358
Qy 358 TYGLAEKILPPAIOISVYVGLPLDOYANSGVFLRWYLIRAAKMELVNVRPEET 417
Db 359 TYGLAEKMLPPAIOISVYVGLPLDOYANSGVFLRWYLIRAAKMELVNVRPEET 418
Qy 418 YRGAIVLEPLRGVHENIAVLDPSSMYPMIKYVNPDTLVPRGKCGEC-GCWEAPEVK 476
Db 419 YRGAIVLEPLRGVHENIAVLDPSSMYPMIKYVNPDTLVPRGKCGEC-GCWEAPEVK 478
Qy* 477 HFFRRCPPGPKVTVLELLELRKRVAAEMKTPPSPPEYRLDROQALKVLANASYGM 536
Db 479 HFFRRCPPGPKVTVLELLELRKRVAAEMKTPPSPPEYRLDROQALKVLANASYGM 538
Qy 537 GMSGARWYCBKAAVTAMGRHLIRTAIIRAKGLKVIYDGTDSLFTVYDPEKVENFKI 596
Db 539 GMSGARWYCBKAAVTAMGRHLIRTAIIRAKGLKVIYDGTDSLFTVYDPEKVENFKI 598
Qy 597 IIKELGPEIKLEKVKYRLLFTEAKKRYAGLLEDRIDIVGFEAVRGDMCELAKEVQTV 656
Db 599 IIKELGPEIKLEKVKYRLLFTEAKKRYAGLLEDRIDIVGFEAVRGDMCELAKEVQTV 658
Qy 657 VEIVLKTSEVNVKAVEYRKIVKELEEGKVPKEKVIYMKTLISKRLSEYTTAPHVAAKRM 716
Db 659 VEIVLKTSEVNVKAVEYRKIVKELEEGKVPKEKVIYMKTLISKRLSEYTTAPHVAAKRM 718
Qy 717 LSAGRVSPGDKIGYIVVKGGRISORAMPYPMVNDPSOLIDVTYVDHOIIPALRIIGY 776
Db 719 LSAGRVSPGDKIGYIVVKGGRISORAMPYPMVNDPSOLIDVTYVDHOIIPALRIIGY 777
Qy 777 FGITEKUKASATGOKTLPDLFLAKK 801
Db 778 FGITEKUKASATGOKTLPDLFLAKK 802

RESULT 3

US-08-062-368-2
Sequence 2, Application US/08062368
Patent No. 5491086
GENERAL INFORMATION:
APPLICANT: Gelfand, David H.
TITLE OF INVENTION: Purified Thermostable Nucleic Acid
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Hoffmann-La Roche Inc.
STREET: 340 Kingsland Street


```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Suite 4700
STATE: Los Angeles
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/688,649
FILING DATE: Filed Herewith
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 220/281
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 776 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-688-649-37

Query Match      29.5%; Score 1237.5; DB 2; Length 776;
Best Local Similarity 36.1%; Pred. No. 2.2e-105;
Matches 301; Conservative 154; Mismatches 283; Indels 95; Gaps 23;

QY 8 VLDSSYEVNGKEPQVLIWGIAENGERVVLIDRSFPYFALLAGADPKQV---AQRIR 63
DB 2 ILDDVYITEBKGPIYRIKFK-KENGFKEIHEHRTFRPIYALLRDSKIEVKITGERHG 60
QY 64 ALSRPKPIIGVEDDKKFKGPRPRVLRIRTVLBEAVREVELYKNDVGEVDLEADIRF 123
DB 61 KIYR---IVDVEKVEKFKLGKPIYWKLYLEHPQDVPTIREKVRHBAVVDIPEYDIPF 116
QY 124 AMRYLIDHDLFFPTWYRVEAEPLKNGKGFYDKYLYVKSREPIYGERLAPTKLPDLRIL 183
DB 117 AKRLIDKGLI-----PNEG-----EELKIL 138
QY 184 AFDIEVYSKQSPRRDPVIVIAKTDGDEVL-----FLAEGDKDEKPIREFVEY 235
DB 139 AFDIETLYHGE-BFGCPITIMISYADENAKVITWKNIDLPYEVVSSEEMIKRLRI 197
QY 236 VKRYPPIIVGYNHHPYLLERARILGIKLDVTRRGALEPTTSVGH---VSVPGR 292
DB 198 IREKDPPIIVYNDSDFPFLARAKELGIKLTIGRD-GSEPFQRIIGDWTAVEVGRI 256
QY 293 NVLDYDAEMPEIKISLEVAEYLGVMKSERVILNMWEIPIYWDPKKRPILLQYAR 352
DB 257 HFDLYHVTITINLPYTLLEVEAL-FGKPEKRYAD--EIAKAMESGELERVATYSM 313
QY 353 DVAATYGLAEKILPFAIOLSYVTGLPLDDQVAMSVGFRLEWYLIARAFAKKEILVPRVE 412
DB 314 EDAKATYELGKPLPMEIQSLRVLGQPLMDVSRSTGNLVEMFLRKAYENENVAAPKPS 373
QY 413 RPE-----ETRYGAILVLEPLRGVHENIAVLDPSSTPNIMIKYKNGPDTLVRGEKCGE 466
DB 374 EEEYORLRBSYTGKPYKEPEKGLMENIVYLDPRALYPSIITITNVPDTLNEG----- 428

```

```

QY 467 CGCMF-APFYKHFRCPCPPEFFKTVLERLLKRAVRAEKKYPDPSPEYRLIDEROAL 525
DB 429 CKXNDYAPGVGHKFCXDIIPGFIPLSLGHLEEROKIKTKKE--TQPIEXKILLDYOKAI 487
QY 526 KVLANSYGYMGSGARWYCRECAKAVTAGRLITTA--INIRKGLKLYIGDTSLEFY 584
DB 488 KLANSPFYGYGAKARWYCKEASBYTAGRKYLILWKELEKGFVLYIDTGLTA 547
QY 585 TYD-----PEKVENFIKILKEELG--FEIKLEKYKRLFTFAKKRYAGLLEDRIDI 635
DB 548 TIGCGSESEIKKALFEVVKYINSKLPGLLELEYEGYKGFV-TKKRAVAIDEEKVIT 606
QY 636 VGFPAVRGDMCELAKEVQTVKVEIVLKTSEVNAKAVEYRKIVLEBEGKPIKLYIMKT 695
DB 607 RGLIEVRDMSEIAKETOARVLETILKHGVEBAVARIVKEVLOKLANYEIPPEKAIYEO 666
QY 696 LSKRLEETTEAPHVVAAKRMLSGYRVSFGDKIGYVIVYGGGRISQRAVPMYFMV--DP 753
DB 667 ITRPLEHYKAIQHVAVAKKLAAGVYKIKPGWVIGIYILRGDGPISNRA--TLAEEYDP 723
QY 754 S--QIDVTYVVDHQLIPALRLIGYEGITEKKLKASATGKTLFDFL-AKSK 803
DB 724 KKAHYDAEYIENQVPAVLRILEGFYRKEDLRVYKTDQVGLTSLWNTIKKSZ 776

RESULT 7
US-07-966-278-1
Sequence 1, Application US/07966278
Patent No. 5489523
GENERAL INFORMATION:
APPLICANT: Mathur, Eric A
TITLE OF INVENTION: EXONUCLEASE-DEFICIENT THERMOSTABLE
TITLE OF INVENTION: PYROCOCUS FURIOSUS DNA POLYMERASE I
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Thomas Fitting
STREET: 12526 High Bluff Road, Suite 300
CITY: San Diego
STATE: CA
COUNTRY: USA
ZIP: 92130
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/966,278
FILING DATE: 19921226
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: STG0133P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-792-3680
TELEFAX: 619-792-8477
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 775 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-07-966-278-1

Query Match      29.5%; Score 1236.5; DB 1; Length 775;
Best Local Similarity 36.2%; Pred. No. 2.7e-105;

```


Wed Apr 23 08:09:10 2003

us-10-034-849-2.rtf

Page 6

Matches	301; Conservative	153; Mismatches	283; Indels	95; Gaps	29
Qy	8	VLDSEYVAVGGEPOVITWIGIAENGERVVLIDSPFPPYLLALPGADPROV----	AORIR	63	
Db	2	ILDVYITBEKKPVIRLF-KENGRFKIEHRTFPPIYALLANDSKIEVKKITGERHG	60		
Qy	64	ALSRKSPILIEVEDDKRYKGRPRRLRTFLPRAVREBELVKNVDVEVLEADIRF	123		
Db	61	KIVR-----IVDEKVEKKFUGKRIYMKYLEHODPPIIREKVRKHPAVNDIFEYDIPF	116		
Qy	124	ANRYLIDHDLFFPFWYREVAEPLENKGRFRDVKYLVKSREPLVGEALAPYLDOLRL	183		
Db	117	AKGYLIDKGL-----E-----WEEB-----BEKIL	138		
Qy	184	APDIEVSKGSPRPEDPVYIVIAVKTDDGEVL-----PIAGDKRKPREEVY	235		
Db	139	ADPITLVHEEB-EGFGPIIMISVDENEAVALYIMKNIDLPIVEVVSEREMIRFPRI	197		
Qy	236	VKRYDPDIIVYNNNNHDPWYLLRRARILGILKIDVTRVGAEPITVSHH-----VSPQL	292		
Db	198	IREKPDILIVYNDSEDFPYLKRAREKGIKLTIGRD-GSEPRMORIGDMTAVEVKRI	256		
Qy	293	ANDLVDAEEMPEKIKSLSEVAEYLVGWKKSERYITNMMEIPDYDDDKRPLLQYAR	352		
Db	257	HPDLVHVITRTINLPYTLAEVVAI-FGKPRKYAD--EIKAMESENLERAKYSM	313		
Qy	353	DOVRATNYGLAKLIPALICLSYVTSGLPIDQCAWSVPRLEWYLLIAAKMELVPRVE	412		
Db	314	EDAKTITELGEPFLPMHIOISRLVQGPLDWSRSTGNLYENMLUKKALYERNEVAPNPS	373		
Qy	413	REP-----ETRYKALYLEPGRVHNIAVLDFSSWPNIMIKYNGPDTLVPRGEKGE	466		
Db	374	BEHYQRIRBESTGFIYAEKBEKLMENIYLDPRLLPSTIIITANVSPTLLBES----	428		
Qy	467	CGGWE-APDYKARFRRCPPGFPTVYLERLLELRKVRPAEMKCTPPDSEYRLIDEROKAL	525		
Db	429	CKNYDAIQVNHKCKOIPGPIPSLLGHLEEROKIKTKQKE--TOPRIKILLDYROAI	487		
Qy	526	KYLANASVYGMWSGARWYCRGCARAVTAMGRHLRTA-INIRAKLGLVIVIGDYSLFV	584		
Db	488	KILANSFYGVYGVAKAMWYCKEABEVTMGRKYELMWKLEEEKGFVLYLIDDTGLYA	547		
Qy	585	TYD-----DEKVENFIKIKIEKLS--PEIKLEKVRKRFPEAKKRVAGLLEBGRDI	635		
Db	548	TIPEGSESEIKKKALLFPVYKINSKLPGLLELEBEGFYKRGFV-TKKRAYIDEKGKIT	606		
Qy	636	VGEPAVRDGMCELAKEVOTKVAIEVLKTSSEVKNANVEYRKIVALEBEGVPREKVIKMT	695		
Db	607	RGLIYVRDMSIAKETQAVLETLIKHDQVBEAVRIYKEVIOQLNANVEIPREKALIEQ	666		
Qy	696	LSKRLLEATTAPVYVNAKRMLSAGTRVSPGKIGYIVVGGGRIGQRMVPYPMVK--DP	753		
Db	667	ITRPLEHKYALGHNAAVAKGLAKGKYLIPGNGIGYILVLRDDGJNNR--TLAEYDP	723		
Qy	754	S--QIDVTYVVDHOLIPALRLIGYGITTEKKLKSATGQKTLPDFL-AKKS	802		
Db	724	KHKHDAEYIENVLPAVRIILIEGGYAKEDLRVQKTRVGGLTSMWLNTKS	775		

RESULT 8
 US-08-424-921-1
 ; Sequence 1, Application US/08424921
 ; Patent No. 5543552
 ;
 ; GENERAL INFORMATION:
 ;
 ; APPLICANT: Mathur, Eric A
 ; TITLE OF INVENTION: PURIFIED THERMOSTABLE PYROCOCCUS
 ; TITLE OF INVENTION: FURIOUS DNA POLYMERASE I
 ; NUMBER OF SEQUENCES: 6
 ;
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: Bingham & Fitting
 ; STREET: 1526 High Bluff Road, Suite 300
 ; CITY: San Diego
 ; STATE: CA

COUNTRY: USA
ZIP: 92130
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/424,921
FILING DATE: 19-APR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/803,627
FILING DATE: 02-DEC-1991
APPLICATION NUMBER: US 07/620,568
FILING DATE: 03-DEC-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/657,073
FILING DATE: 19-FEB-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/776,552
FILING DATE: 15-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: STG0100P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-792-3680
TELEFAX: 619-792-8477
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 775 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
ANTI-SENSE: NO
US-08-424-921-1

```

Query_Match Similarity      29.5%, Score 1236.5; DB 1, Length 775;
Best Local Similarity      36.2%, Pred. No. 2,7e-105;
Matches 301; Conservative 153; Mismatch 283; Indels 95; Gaps 23

QY      8 VLDSEYVNGKEPQVITWGIENGERVVLIDRSRPFYALLAPGADPKQV---AQRIR 63
DB      2 ILDVVITIEGKRPPIRLFK-KENGGFKETIENHRTFRPIYALLDSDSKIEEYKKTIGRHG 60
QY      64 ALSRPKSDIIGVEDDKKATGFGRRPRVRLARKTLEAVREYRELKYNVDGVEDVLEADIR 123
DB      61 KIVR-----IVAVEKVEKKELGKRIITWKTLYLEHPODPIIRREVRHNPAVVDIEFYDIP 116
QY      124 AKRYLIDHDLPPFTWYVREAEPLENNMGFRVDKYYLKVSRREPLVEALAPTKPDLRI 183
DB      117 AKRYLIDGILL-----PMEGE-----EELKIL 138
QY      184 AEDIEVSKQSPREPERDPIYIAVKTDDGEVL-----PIAGKODRKPIREFVEY 235
DB      139 APDIEIYLHEGE-BFGKGPITMSYADENNAKAVITWKNIDLPVLEVVSERMIKRFRLI 197
QY      236 VKRYQPDITVYNNNNHMDWYLLRPARILGIKIDVTRVGAFTSYNGH---YVSQGR 252
DB      198 IRRKQPDITVYNNDSDFPYLAKGAELKGIKLTIGND-GSEKQKORIGDMTAYEVGRI 256
QY      293 NVDLIDVAEEMPEIKISLEEAELVGVKKKSEVITINMVELIDYVDDEKKRPILLQYAR 352
DB      257 HEPFLVHIVTRITINPIYTLLEAYEAI-GKKREKRVYAD-EIKAAESENLEERAAKYSM 313
QY      353 DVPVATYGLAEKILPFAIQLSYVTGLPDDQVGAHSVGRLEWYLLPAAFKKELVPRVE 412
DB      314 EDAKATYELGKEPLPMEIQLSRVQOPLMDSVRSSTNVLWEMFLILARKYERNEVAPKPS 373
QY      413 RPE-----EYRAIVALEPLRGCHENIAVLDFSMTPNIMIKYVNGPDLTVRQGEKGE 466

```


Db 374 EEEYORRLRESYTGPFVKEPEKGLMENIVYLDPRALYPSIIITHNVSPDTLNEG----- 428
 Qy 467 CGCME-APEVGRFRRCPPGFVKTVLELLLELRKRVAEEMKKYPPDSEYLLDEROKAL 525
 Db 429 CKNYDIAPOVGHKFCODIPGIFPSLGHLEBEROKITKMBE-TODPIEKLILYDROKAI 487
 Qy 526 KVLNANASYGMGSGARWYCEKACAVTAMGRHLIRTA-INIARKLGKLYIGDTSLFV 584
 Db 488 KLANSPFYGYGAKARWYCEKACASVTAMGRKXIELVWKELEKFGKLYIDTDLGYA 547
 Qy 585 TYD-----PEKYENFIKIIKEELG--FEIKLEKYKRLFFTEAKKRYAGLLEDGRIDI 635
 Db 548 TIPGSESEIEKKALFEFKYINSKLPGLLELEEGFYKRGFFV-TKKRAYVIDEGKVIIT 606
 Qy 636 VGEFAVAGDMCELAKEVOTKVEIYVLTSEVNKAVEYVRKI VKELEBKRYIEGLVIMKT 695
 Db 607 RGLIIVRDMSEIAKETQAVLETILKHGDVEAVRIVKEVIOKLANEIPPEKLAIEQ 666
 Qy 696 LSKRLEETTEAPHVAAKRLMISAGYVSPGDKIGYIVKGGRIISORAMPYFMVK--DP 753
 Db 667 ITRPLHEKXKAGPHVAVAKKLAAGVKIKRGMVIGIYVLRDGPISNRA---ILAEYDP 723
 Qy 754 S--QIDVTYVDHQIIPALRIIGYIGITEKKLASATGQTLFDFL-AKKS 802
 Db 724 KKHKYDAEYIENQVLPVAVIRILEGFGYRKEDLRYOKTROVGLTSMINIKKS 775

RESULT 9
 US-08-556-355A-1
 ; Sequence 1, Application US/08556355A
 ; Patent No. 5866395
 ; GENERAL INFORMATION:
 ; APPLICANT: MATTHEW, Eric J.
 ; TITLE OF INVENTION: Purified Thermococcal Pyrococcus
 ; TITLE OF INVENTION: Purified Thermococcal Pyrococcus I
 ; NUMBER OF SEQUENCES: 6
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Flanagan, Henderson, Farbow, Garrett &
 ; STREET: 1300 I Street, N.W.
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20005
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/556,355A
 ; FILING DATE: 13-NOV-1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/424,921
 ; FILING DATE: 19-APR-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/803,627
 ; FILING DATE: 02-DEC-1991
 ; APPLICATION NUMBER: US 07/779,846
 ; FILING DATE: 21-OCT-1991
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/657,073
 ; FILING DATE: 19-FEB-1991
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/620,568
 ; FILING DATE: 03-DEC-1990
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Barker, M. Paul
 ; REGISTRATION NUMBER: 32,013
 ; REFERENCE/DOCKET NUMBER: 04121.0004-02
 ; TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)408-4000
 TELEFAX: (202)408-4400
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 775 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-556-355A-1

Query Match 29.5%; Score 1236.5; DB 2; Length 775;
 Best Local Similarity 36.2%; Pred. No. 2,7e-105;
 Matches 301; Conservative 153; Mismatches 283; Indels 95; Gaps 23;

Qy 8 VLDSYEVGKEPQVLIWGIAENGERRVLLIDRSFRPYFALLAPGADPKQV---AQRIR 63
 Db 2 ILVDVYTEBGKPIYIRLFK-KENQKFIKHDRTRFRPIYALLDSDKIEEVKKTIGERHG 60
 Qy 64 ALSRPKPIIGVEDDKKTKYRPRVRIRITVLEAVREYRELKQVNDVEDVLEADIRF 123
 Db 61 KIVR---IVDVEKVEKKFKGPIYWKLYLEHPODPTIREKVRHAPAVDIFEYDIPF 116
 Qy 124 AMRYLIDHDFPFTWYRVAEPLNKKGFVADVYLVKSRPEPLYGKALAPTKLPDLRI 183
 Db 117 AKRYLIDKGLI-----PDEG-----BELKIL 138
 Qy 184 AFDEIVYSKQSPRPERRPVIVIAVKTDDGEVL-----FIAEGKDRKPIREFVEY 235
 Db 139 AFDIETLYHGE--BFGKPIIMISYADENKAVITWKNIDLPYVEVSSREMIKRLRI 197
 Qy 236 VKRYPPIIYGNNNHDMYLLRRARILGKIDVTRRVAEPTTSVGH---VSPGRL 292
 Db 198 IREKDPPIIYVNGDSDFPYLAKRAEKLGIKLTIGRD-GSEPMORIGMTAVEVGR 256
 Qy 293 NVLDYDAEEMPEIKIKISLEVAEYGVKSKSERVIINMEIPDYDDPKRRPLLQYAR 352
 Db 257 HFDLYHYITRTINLPYTLAEVTAI-FGKPKREYVAD--EIKAMSGENLERVAKYSM 313
 Qy 353 DDAVATYGLAEKILPFAIQLSYVTGLPLDOVGAMSVGFLEWYILRAFKELVPRVE 412
 Db 314 EDKATYELKEPFLPMGIQISRLVCGPLMDVSRSTGNLVEMFLRRAVERNEVAPRKS 373
 Qy 413 RPE-----ETRGALVLEPLRGVHENIAVLDPSMTDNIMIKYNGPDTLVPRGEKGE 466
 Db 374 EEEYORRLRESYTGPFVKEPEKGLMENIVYLDPRALYPSIIITHNVSPDTLNEG----- 428
 Qy 467 CGCME-APEVGRFRRCPPGFVKTVLELLLELRKRVAEEMKKYPPDSEYLLDEROKAL 525
 Db 429 CKNYDIAPOVGHKFCODIPGIFPSLGHLEBEROKITKMBE-TODPIEKLILYDROKAI 487
 Qy 526 KVLNANASYGMGSGARWYCEKACAVTAMGRHLIRTA-INIARKLGKLYIGDTSLFV 584
 Db 488 KLANSPFYGYGAKARWYCEKACASVTAMGRKXIELVWKELEKFGKLYIDTDLGYA 547
 Qy 585 TYD-----PEKYENFIKIIKEELG--FEIKLEKYKRLFFTEAKKRYAGLLEDGRIDI 635
 Db 548 TIPGSESEIEKKALFEFKYINSKLPGLLELEEGFYKRGFFV-TKKRAYVIDEGKVIIT 606
 Qy 636 VGEFAVAGDMCELAKEVOTKVEIYVLTSEVNKAVEYVRKI VKELEBKRYIEGLVIMKT 695
 Db 607 RGLIIVRDMSEIAKETQAVLETILKHGDVEAVRIVKEVIOKLANEIPPEKLAIEQ 666
 Qy 696 LSKRLEETTEAPHVAAKRLMISAGYVSPGDKIGYIVKGGRIISORAMPYFMVK--DP 753
 Db 667 ITRPLHEKXKAGPHVAVAKKLAAGVKIKRGMVIGIYVLRDGPISNRA---ILAEYDP 723
 Qy 754 S--QIDVTYVDHQIIPALRIIGYIGITEKKLASATGQTLFDFL-AKKS 802
 Db 724 KKHKYDAEYIENQVLPVAVIRILEGFGYRKEDLRYOKTROVGLTSMINIKKS 775

RESULT 10

Wed Apr 23 08:09:10 2003

us-10-034-849-2.ra1

Page 8

US-07-803-627A-1
Sequence 1, Application US/07803627A
Patent No. 5948663
GENERAL INFORMATION:
APPLICANT: MATHUR, Eric J.
TITLE OF INVENTION: Purified Thermotable Pyrococcus
TITLE OF INVENTION: Purified Thermotable Pyrococcus
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finegan, Henderson, Farabow, Garrett &
ADDRESSEE: Dunner, L.L.P.
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/803,627A
FILING DATE: 02-DEC-1991
CLASSIFICATION: 415
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/776,552
FILING DATE: 14-OCT-1991
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/657,073
FILING DATE: 19-FEB-1991
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/620,568
FILING DATE: 03-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: Barker, M. Paul
REGISTRATION NUMBER: 32,013
REFERENCE/DOCKET NUMBER: 04121,0004-00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)408-4000
TELEFAX: (202)408-4000
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 775 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
-07-803-627A-1
Query Match 29.5%; Score 1236.5; DB 2; Length 775;
Best Local Similarity 36.2%; Pred. No. 2.7e-105;
Matches 301; Conservative 153; Mismatches 283; Indels 95; Gaps 23;
QY 8 VLDSSYEVVKGKPOVITWIGIAENGERVVLIDSPFYPYALAGADPKOV---AQRIR 63
DB 2 ILVDVITEEGKPVRLFK-KENGKFKIEHDFRPRVYIALLRDSDKIEVKKITERBG 60
QY 64 ALSRPSPFIVGDDKRYGPRPRVLRIRTVLPEAVREVELVANDGVEDLVADIRF 123
DB 61 KIVR-----IVDEKYEKKFKLQKPTWKLYLHHPQOVPIREKREHRAVVDIFEDIRF 116
QY 124 ARVYLIDHDLFFFTVYRREAPLENKKGFRVVDVYLVSRPEPLXGELAPTLPLRL 183
DB 117 AKRYLIDKGLI-----PMEGG-----EELKRL 138
QY 184 APDIEVYSKQSPREPRDVIVAKTDGDEVL-----FLAEGKDRKREPRFVY 235
DB 139 APDIEVYSKQSPREPRDVIVAKTDGDEVL-----FLAEGKDRKREPRFVY 235
QY 236 VKRYPDIIIVGNNHFWPVLARAILGIKLDVTRVGAEPFTSVHGH---VSPGRL 292
DB 198 IREKPDIIIVTNGDSDFPYLAKRAKGLIKLTIGRD-GSPKMGRIGMTAVVEVKRI 256

QY 293 NVLDVYAEEMPEIKISLEVEVYLVGMKSRVYIIMMELPDVDDPKRPLLLQYAR 352
DB 257 HFDLYVITRTINLTPTYLEAVYAL-FGRKPKRYAD-ELAKMSENGENLERYAKYGM 313
QY 353 DVPATYGAELKILPFAIQLSYVTGLPLDOYGMVSGRELEWYLIRAFKKELVNRVE 412
DB 314 EDKATYELGKELPMEIQLSRLVGOPLDVSRSSSTGNLVEFLRKAYERNSEAPAKKS 373
QY 413 RPE-----ETVRAIVLEPLRGVHENIAVLDSSNPNYIMIKYVGPDTLVPRGKCE 466
DB 374 EEEYQRLRESYTGFGVPEPEKGLMEYIVLFRLLYRSIIITHNVSBDTLNEG----- 428
QY 467 CGCME-APVKRFRCPGPFKTVLERLELRKVRLEMKKYPPDSPEYRLIDEROKAL 525
DB 429 CKRYDIAPOVGHKFCOKIPGFIPLGLHLLEEROKIKTMEKE-TODPLEKILLDYRQKA 487
QY 526 KVLANSYGMGSGARWYCECAKAVTAMGRHLIRTA-INIARRGLKTYGDTSLPV 584
DB 488 KLANSPFYGYVAKARWYCECAKAVTAMGRKYLELWKELEKRFKVLITDTGGLVA 547
QY 585 TYD-----PEKVENFKIKIEELG--PEIKLEKYKLEFTEAKRYAGLEDERDOI 635
DB 548 TFGGSEELKKGALFEKYNSTLPGLELEFGYRKGFPV-TKRYAVIIDEKGVLT 606
QY 636 VGFPAVRCWCELANEVOTKVEIIVLKTSEVNKAVERKIVKLEEGKVPLEKVIWKT 695
DB 607 RGLIVRRMSIAKETARVLETILKHGDVEAVRIVEYIQLKLANYEIIPKRLAIYEQ 666
QY 696 LSRLEETTEAPHVAAKRLSNAGYVSPDGIGVIVKGGRIQRAPYEMVK--DP 753
DB 667 ITRPLEHYKAIQPHVAVAKKILAAKGVKIKKGMWIGYIVLRGDSISRA---ILAEYDP 723
QY 754 S--QIVVTYVHQIIPALRLILGFGYTEKLLKASATGQKTLDFPL-AKKS 802
DB 724 KKKYDAEYIENQVLPALRLILGFGYTEKLLKASATGQKTLDFPL-AKKS 775
RESULT 11
US-09-073-354-1
Sequence 1, Application US/09073354
Patent No. 6038659
GENERAL INFORMATION:
APPLICANT: KITABAYASHI, Masao
APPLICANT: ARAKAWA, Taku
APPLICANT: INOUE, Hiroaki
APPLICANT: KAMAKAMI, Bunsei
APPLICANT: KAMAMURA, Yoshihisa
APPLICANT: IMANAKA, Tadayuki
APPLICANT: MORIKAWA, Masaaki
TITLE OF INVENTION: A Thermotable DNA Polymerase and Kits for
TITLE OF INVENTION: Amplifying Nucleic Acids
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kenyon & Kenyon
STREET: 1025 Connecticut Avenue, N.W., Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 6.1 Windows
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/073,354
FILING DATE: Concurrent Herewith
CLASSIFICATION: 0506
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/656,005
FILING DATE: 24 MAY 1996

APPLICATION NUMBER: JP 134096/95
 FILING DATE: 31 MAY 1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Toffenetti, Judith L.
 REGISTRATION NUMBER: 39,048
 REFERENCE/DOCKET NUMBER: 2418/9
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-429-1776
 TELEFAX: 202-429-0796
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 774 amino acids
 TYPE: amino acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-073-354-1

Query Match 29.5%; Score 1234.5; DB 3; Length 774;
 Best Local Similarity 36.4%; Pred. No. 4,2e-105;
 Matches 302; Conservative 145; Mismatches 289; Indels 93; Gaps 22;

QY 8 VLDSSYEVNGKEPOVITWGAENGERVVLIDRSFRPFYALLAPGADPKOV---AQRIR 63
 2 ILTDYITDEDKPVIRIFK-KENGSEFKIEYDRTEPEFYALLKDDSAIEVKKITAERHG 60
 DB 64 ALSRPKPIIGVEDDKRYGRPRVLRIRTVLPEAVREYRELKVDGVEDVLEADIRF 123
 61 TVVTVKR---VEKVQKFKGRPEVWKLFTTHPDVPAIRDKIREHGAVIDIYEYDIPF 116
 QY 124 AMRYLIDHDLFPFTWYVEAEPLNKGFRVDKYLVSREPELYGEALAPTKLPDLRL 183
 117 AKRLIDKGLV-----PMEGD-----ELKML 138
 QY 184 AFDIEVYSGKSGSPRPDPVIVIAVKTDGSEVL-----FLAEGKDRKPIREPEVEY 235
 139 AFDIQTLYHEGESEPAE-GRIMISYADEEGARVITWKKVLD.PYVDVSTEREMIKRFLRV 197
 QY 236 VKRDPDIIIVGNNHNPWPLRLRRLIGTLDVTRVNGEPTTSVGH---VSVPRL 292
 198 VKERDPDLITVNDNPFALVKRCCEKLGINFALGRD-GSEPKIORMGDRFAVEVKRI 256
 QY 293 NVLDYVAEEMPEIKIKSLSEVAEYLVGKKSERVIINWMEIIPYWDPKRPLLOYAR 352
 257 HFDLYPVIRITINPTTLEAVENAVGQPK-EKYVAE-ELTPAMEGLENLERYAKISM 313
 QY 353 DDVATYGLAEKILPFAIOLSYVTGLPLDOYGANSVGRLEMYLIRAAFKMKELVPPNVE 412
 314 EDAKVTYELGKEPLPMEAQLSRLIGQSLMDVSRSTGNLVEMPLLRKAYENNELAPNKP 373
 413 RPE-----ETRGALVLEPLAGHENLAVDPSSMYPMIMKINVGPTLVPRGKCGEC 467
 374 EKEIARRRQSEYGVKEPERGLMENVYLDLFRSLYPSIIITHNVSPTLNRE----- 426
 QY 468 GCWE---APEVGRFRRCPPGFPKTVLERLLELRKVAEMKKCPDPSREYRLDEROKA 524
 427 GCKEYDVAPOVGHFCDPFGFISLDLLEBKOKIKKKA-TIDIEKKLADYRRA 485
 QY 525 LKVLANSYGYMGWSGARWYCRECAVTAAGRHLIRTAI-NIARKLGLKVIYGDPTSLF 583
 486 IKILANSYGYGYVARRAWYCKECAESVTANGREYITWTIKIEBKQYFKVIYSDTDF 545
 QY 584 VTY---DPEKE---NFIKIKEEL-GFEIKLEKYKLFTEAKKRYAGLLEDRID 634
 546 ATIPGADAEIVYKKAMEFLINAKLPALGALIEYEGYKRGFFV-TKKKVIADIEBKIT 604
 QY 635 IVGEAVNGDWCELAKEVQTVKVEIVLKTSEYNKAVEYVRKIVLELSEKXPIELGLVWK 694
 605 TRGLEIYRRDSEIATKEQAVLEALLKDGVEKAVRIVKAYTEKLSKYEVBPPEGLVHE 664
 QY 695 TLSRLIEEYTTAPHVAAKMLAGSVSPQDKIGYIVKGGGRISORAMPYFWVNDPS 754
 665 QITBDLKDYKATGPHVAVAKRLAAGVIRPGTVISYIVLKGSGRIGRALP-FDEFPPT 723

QY 755 --QIDVTYVVDHOLIPALRIILGFITTEKULKASATGKTLFDPLAKK 801
 DB 724 KHKYDAEYIENOVLPAPERILRAPFGRKEDLRYQKTRQVGLSAMLKP 772

RESULT 12
 US-08-656-005A-1
 Sequence 1, Application US/08656005A
 Patent No. 6054301
 GENERAL INFORMATION:
 APPLICANT: KITABAYASHI, Masao
 APPLICANT: ARAKAWA, Taku
 APPLICANT: INOUE, Hiroaki
 APPLICANT: KAWAKAMI, Bunsei
 APPLICANT: KAWAMURA, Yoshihisa
 APPLICANT: IMANAKA, Tadayuki
 APPLICANT: TAKAGI, Masahiro
 APPLICANT: MORIKAWA, Masaaki
 TITLE OF INVENTION: A Method of Amplifying Nucleic
 TITLE OF INVENTION: Acid and A Reagent Therefor
 NUMBER OF SEQUENCES: 16
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Kenyon & Kenyon
 STREET: 1025 Connecticut Avenue, N.W., Suite 600
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20036
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Wordperfect 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/656,005A
 FILING DATE: 24 May 1996
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 134096/95
 FILING DATE: 31 May 1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Toffenetti, Judith L.
 REGISTRATION NUMBER: 39,048
 REFERENCE/DOCKET NUMBER: 2418/3
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-429-1776
 TELEFAX: 202-429-0796
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 774 amino acids
 TYPE: amino acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-656-005A-1

Query Match 29.5%; Score 1234.5; DB 3; Length 774;
 Best Local Similarity 36.4%; Pred. No. 4,2e-105;
 Matches 302; Conservative 145; Mismatches 289; Indels 93; Gaps 22;

QY 8 VLDSSYEVNGKEPOVITWGAENGERVVLIDRSFRPFYALLAPGADPKOV---AQRIR 63
 2 ILTDYITDEDKPVIRIFK-KENGSEFKIEYDRTEPEFYALLKDDSAIEVKKITAERHG 60
 DB 64 ALSRPKPIIGVEDDKRYGRPRVLRIRTVLPEAVREYRELKVDGVEDVLEADIRF 123
 61 TVVTVKR---VEKVQKFKGRPEVWKLFTTHPDVPAIRDKIREHGAVIDIYEYDIPF 116
 QY 124 AMRYLIDHDLFPFTWYVEAEPLNKGFRVDKYLVSREPELYGEALAPTKLPDLRL 183
 117 AKRLIDKGLV-----PMEGD-----ELKML 138

Wed Apr 23 08:09:10 2003

us-10-034-849-2.rst

Page 10

QY 184 AFDIEVYSKQSPRRERDPIVIAVKTDDGEVL-----FLAEKDRKPIREFEVY 235
DB 139 AFDIOTLYHEGEFAE--GPILMISYDEBEGARVITWKNVDLPYDVVSTREMIKRFLLV 197
QY 236 VKRYDPDIIIVGYNHNPMPYLLRARIIGIKLDVTRVGAEPPTVYGH---VSYPGRU 292
DB 198 VKKEDPDVITVNGDNDFALYKRCCKGGINFALGRD--GSEPKIORMGROFVAVYKRI 256
QY 293 NVDLVYAEEMPEIKISLEVAEYLGVMKSSRVYIMMWEIPYMDPKKRLLOYAR 352
DB 257 HFDLYPVIRRTINLPYTLLEAVYAVGQPK--EKVYAE--ETPAMETGENLBRVARYGM 313
QY 353 DVVRATYGLAEKILFPALQLSYVTGLPLDOYGAMSVPFLLEYILRAAFKMKELVNRVE 412
DB 314 EDKATYTELKGEFLPMKQSLRSLGOSLMDVSSSTGNLVEMFLRKATYERNEBLAPKPD 373
QY 413 RPE-----ETRGATVLEPLRGVHENIIVLDFSSMYPNIMIKYVGPDTLVPRGCKGEC 467
374 EKEIARRROSYYEGYVKEPERGEMENIVYLDPSILPSTIITHNVSPDTLNRE----- 426
468 GCWE--ABEVKGRFRRCPPGFVTVLERLLERKRVRAEMKYPSPSPRYLLDEROKA 524
DB 427 GCKEYDVAPQVGHRCXDPGPFIPSLIGDLLEBRQIKKKMKA--TIDPIERKLLDYROBA 485
QY 525 LKVLNASTGYMGSGARWYCRECAKAVTAMGRHLRPAI--NIARKGLKVIYGDPTSLF 583
DB 486 IKILANSYGYGYVABARWYCKECSAVTAMGREYITWTIKIEBKYGPFVITSDTGF 545
QY 584 VTY---DPEKVE---NFIKIKKEU--GPEIKLEKVVYRLPFTBAKKRYVAGLLEGRD 634
DB 546 ATIPGADAEIVYKKAEMFLNINAKLPGALELEBEFYKGFV--TKKTAIVADERKLT 604
QY 635 IVGEFVARDGCELAKEVQTVNVEIYVKTSEVNAKAVEYRKIVKLEBGKVPLEKVIYK 694
DB 605 TRGLEIVRDMSEIARQAVLALBKODQVKAIRIVKETEKLSTYEPPEKVIYHE 664
QY 695 TLSKRLSEYTTAPVNAKMSAGYRVSPGDKIGYIVVKGGRISORAMPYFMVKDS 754
DB 665 QITRDKDKYATGPVNAVAKKLAARVYKIRPGTVISYVLKSGSGIDGRAP--FDFEDPT 723
QY 755 --QIDVTYVVDHIIIPALRIILGYFITEKKLKASATGOKLFPFLAK 801
DB 724 KHKYDAEYIENQVLPAYERILRAFGRKEDLRYQTRQVGLSAMLKPK 772

RESULT 13

US-09-073-259-1

Sequence 1, Application US/09073259

Patent No 6143536

GENERAL INFORMATION:

APPLICANT: TANAKA, Tadayuki

APPLICANT: TAKAGI, Masahito

APPLICANT: MORIKAWA, Masaki

TITLE OF INVENTION: DNA encoding a Thermocable DNA Polymerase

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: Kenyon & Kenyon

STREET: 1025 Connecticut Avenue, N.W., Suite 600

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20036

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WordPerfect 6.1 Windows

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/073,259

FILING DATE: Concurrent Herewith

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/656,005

FILED DATE: 24 MAY 1996
APPLICATION NUMBER: JP 134096/95
FILING DATE: 31 MAY 1995
ATTORNEY/AGENT INFORMATION:
NAME: Toffenetti, Judith L.
REGISTRATION NUMBER: 39,048
REFERENCE/DOCKET NUMBER: 2418/10
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-429-1776
TELEFAX: 202-429-0796
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 774 amino acids
TYPE: amino acid
STRANDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-073-259-1

Query Match 29.5%; Score 1234.5; DB 4; Length 774;
Best Local Similarity 36.4%; Pred. No. 4,2e-105;
Matches 302; Conservative 145; Mismatches 289; Indels 93; Gaps 22;

QY 8 VLDSYEVVGEKPEVYIINGIAENGERVVLIDRSFRPYFALLAGADPKOV---AORIR 63
DB 2 ILDTDVTEDGKPVIRIK--KENGEPIEBYDRFEPYFALLKODSAIEVYKKTITERHG 60
QY 64 ALSRPSPILIGVEDDKRYRGRPRVLRIRTVLPKAYREYELVKNVDGVEDLLEADIRF 123
DB 61 TVTVYKR---VEKQKFLGRPYEVWKLFTTHPDVPAIRKIRHGNVYDIYETDIRF 116
QY 124 AMRYLIDHOLFPTTVRYVAEPLLENKGRFVDKVIKVSREPLYGELAPYLPLRL 183
DB 117 AKRLLIDKGLV-----PMEGD-----EELKML 138
QY 184 AFDIEVYSKQSPRRERDPIVIAVKTDDGEVL-----FLAEKDRKPIREFEVY 235
DB 139 AFDIOTLYHEGEFAE--GPILMISYDEBEGARVITWKNVDLPYDVVSTREMIKRFLLV 197
QY 236 VKRYDPDIIIVGYNHNPMPYLLRARIIGIKLDVTRVGAEPPTVYGH---VSYPGRU 292
DB 198 VKKEDPDVITVNGDNDFALYKRCCKGGINFALGRD--GSEPKIORMGROFVAVYKRI 256
QY 293 NVDLVYAEEMPEIKISLEVAEYLGVMKSSRVYIMMWEIPYMDPKKRLLOYAR 352
DB 257 HFDLYPVIRRTINLPYTLLEAVYAVGQPK--EKVYAE--ETPAMETGENLBRVARYGM 313
QY 353 DVVRATYGLAEKILFPALQLSYVTGLPLDOYGAMSVPFLLEYILRAAFKMKELVNRVE 412
DB 314 EDKATYTELKGEFLPMKQSLRSLGOSLMDVSSSTGNLVEMFLRKATYERNEBLAPKPD 373
QY 413 RPE-----ETRGATVLEPLRGVHENIIVLDFSSMYPNIMIKYVGPDTLVPRGCKGEC 467
374 EKEIARRROSYYEGYVKEPERGEMENIVYLDPSILPSTIITHNVSPDTLNRE----- 426
468 GCWE--ABEVKGRFRRCPPGFVTVLERLLERKRVRAEMKYPSPSPRYLLDEROKA 524
DB 427 GCKEYDVAPQVGHRCXDPGPFIPSLIGDLLEBRQIKKKMKA--TIDPIERKLLDYROBA 485
QY 525 LKVLNASTGYMGSGARWYCRECAKAVTAMGRHLRPAI--NIARKGLKVIYGDPTSLF 583
DB 486 IKILANSYGYGYVABARWYCKECSAVTAMGREYITWTIKIEBKYGPFVITSDTGF 545
QY 584 VTY---DPEKVE---NFIKIKKEU--GPEIKLEKVVYRLPFTBAKKRYVAGLLEGRD 634
DB 546 ATIPGADAEIVYKKAEMFLNINAKLPGALELEBEFYKGFV--TKKTAIVADERKLT 604
QY 635 IVGEFVARDGCELAKEVQTVNVEIYVKTSEVNAKAVEYRKIVKLEBGKVPLEKVIYK 694
DB 605 TRGLEIVRDMSEIARQAVLALBKODQVKAIRIVKETEKLSTYEPPEKVIYHE 664
QY 695 TLSKRLSEYTTAPVNAKMSAGYRVSPGDKIGYIVVKGGRISORAMPYFMVKDS 754

Db 665 QITRDLKDYKATGPHVAARAKLAAGVIRGTVISYVLKSGRIGDRAIP-FDEFDP 723

Qy 755 --QIDVTYVVDHQLIPALALRILGFGITEKKLKSANQKTLFDFLAKK 801

Db 724 KHKYDAEYIENQVLPAPERILRAFGRKEDLRYOKTRQVLSAMLKPK 772

RESULT 14

US-09-363-095-1

Sequence 1, Application US/09363095

Patent No. 6187573

GENERAL INFORMATION:

APPLICANT: IMANAKA, Tadayuki

APPLICANT: TAKAGI, Masahiro

APPLICANT: MORIKAWA, Masaaki

TITLE OF INVENTION: DNA Encoding A Thermostable DNA Polymerase

FILE REFERENCE: 2418/11

CURRENT APPLICATION NUMBER: US/09/363,095

CURRENT FILING DATE: 1999-07-30

EARLIER APPLICATION NUMBER: US 09/073,259

EARLIER FILING DATE: 1998-05-06

EARLIER APPLICATION NUMBER: US 08/656,005

EARLIER FILING DATE: 1996-05-24

EARLIER APPLICATION NUMBER: JP 134096/95

EARLIER FILING DATE: 1995-05-31

NUMBER OF SEQ ID NOS: 16

SOFTWARE: WordPerfect 6.1 windows

SEQ ID NO 1

LENGTH: 774

TYPE: PRT

ORGANISM: Hyperthermophilic archaeon

US-09-363-095-1

Query Match 29.5%; Score 1234.5; DB 4; Length 774;

Best Local Similarity 36.4%; Pred. No. 4.2e-105;

Matches 302; Conservative 145; Mismatches 289; Indels 93; Gaps 22;

Qy 8 VLDSEYVWGKPEPOVITMGANGERVYLIDRSFRPYFALALGADPKQY----AQRIR 63

Db 2 ILDDYITDECKVIRIRFK-KENGEFKIEYDRTPEYFALKDSDALEYKKTAEHSG 60

Qy 64 ALSRPKSPILGVEDDKRYGPRPRVLRIRTVLPEAVREYRELKVNQVDEVDLEADIR 123

Db 61 TVTVVTKR---VEKVKCKFGIRPVEWMLYPTHQDVPAIDKIREHNAVIDIYEIDPF 116

Qy 124 AMRLIHDHLEPFTWYRVEAPELENKGFVQDKVYLKSRPEPLYGEALAPTKLPDLRIL 183

Db 117 AKRYLIDKGLV-----PMEGD-----EELKML 138

Qy 184 AFDIEVYKSGSPRPERDPIVIAVKTDGDEVL-----FLAEGDDKRPREFEY 235

Db 139 AFDIQTLYHGESEFPAE-GPILMISYADSEGAVITKXNDLPYDVVSTEREMIKRFLRV 197

Qy 236 VKRYDPDIIVGNNNNHFWPILRRARILGKLDVTRVGALEPTTSVGH---VSVPGL 292

Db 198 VKKQDPVLITVYNDNDFPALKKRCGLGINPALGRD-GSEPIKQRMGDFPAVEVKGRI 256

Qy 293 NVDLVDAEMPEIKISLEAEVAVLVGKKSERVILINMWEIPIYWDPKKRPLLOYAR 352

Db 257 HFDLYPIRRTINLPYTLAEVAVFQPK-EKYVAE--EITPAMEGENMLERVARYSM 313

Qy 353 DDVATGIALEKILPFAIOQSVTVGLPLDOYGAMSVGRLEWYILIRAFKKEVLPRNVE 412

Db 314 EDAAVTELGKEFLPMEAQSLRGLGOSIMDVSRSTGLVWVFLLRKAYENNELAPNPD 373

Qy 413 RPE-----ETRYGAILVEPLRGVHENIIVLDFSSMYPNIMIKYVGPPTLVPRGKSEC 467

Db 374 EKELARRRQSYEGGVKEPREGIMENIYVLDRLSLPESIIITTHVSPPTLARE----- 426

Qy 468 GCWE---APEVKARFRCPGPFKTVERLLLELRKRVAEKKYPPDSPEYRLDEROKA 524

Db 427 GCKEYDVAPOGVHRFCDFPGFISLLGLDLLEEROKIKKKMKA-TIDIERKLLDYRQRA 485

Qy 525 LKVLNANSGYMGWMSGARWYCRECAKAVTAMGRHLIRTAI-NIARKLGLKVIYGDTSLF 583

Db 486 IKTLANSYGYIGYARARAKWCKEASVTAMGREYITMTKEIBEKGFVITSDTGGF 545

Qy 584 VTY---DPEKVE---NFIKIIEEL--GFEIKLEKYYKLFTEAKKRYAGLLEDGRID 634

Db 546 ATIPGDAEIVKKKAMEFLVYINAKLPGLALEYEGFYKRGFFV-TRKKYAVIDEBGKIT 604

Qy 635 IVGFVAVRGWCLAEVQTRVVEIVLKTSEVNKAVAYRYKIKYKELEBEGVPIEKLYWK 694

Db 605 TRGLEIVRRDMSIETAKETQARVLEALLKQDVKEKAVIYKEVTEKLSKYEVPEKLVH 664

Qy 695 TLSKRLSEYTTAAPHVAAAKRMLSAGRYVSPGKIGVVIYKGGRIISQRAWPFVMDPS 754

Db 665 QITRDLKDYKATGPHVAARAKLAAGVIRGTVISYVLKSGRIGDRAIP-FDEFDP 723

Qy 755 --QIDVTYVVDHQLIPALALRILGFGITEKKLKSANQKTLFDFLAKK 801

Db 724 KHKYDAEYIENQVLPAPERILRAFGRKEDLRYOKTRQVLSAMLKPK 772

RESULT 15

US-09-418-027-1

Sequence 1, Application US/09418027

Patent No. 6225065

GENERAL INFORMATION:

APPLICANT: KITABAYASHI, Masao

APPLICANT: ARAKAWA, Taku

APPLICANT: INOUE, Hiroaki

APPLICANT: KAWAKAMI, Bunsei

APPLICANT: KAWAMURA, Yoshihisa

APPLICANT: IMANAKA, Tadayuki

APPLICANT: TAKAGI, Masahiro

APPLICANT: MORIKAWA, Masaaki

TITLE OF INVENTION: A Thermostable DNA Polymerase and Kits for

TITLE OF INVENTION: Amplifying Nucleic Acids

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: Kenyon & Kenyon

STREET: 1025 Connecticut Avenue, N.W., Suite 600

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20036

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WordPerfect 6.1 windows

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/418,027

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/073,354

FILING DATE:

APPLICATION NUMBER: JP 134096/95

FILING DATE: 31 MAY 1995

ATTORNEY/AGENT INFORMATION:

NAME: Tofenetti, Judith L.

REGISTRATION NUMBER: 39,048

REFERENCE/DOCKET NUMBER: 2418/9

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-429-1776

TELEFAX: 202-429-0796

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 774 amino acids

TYPE: amino acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-418-027-1

Wed Apr 23 08:09:10 2003

us-10-034-849-2.raf

Page 12

Query Match 29.5%; Score 1234.5; DB 4; Length 774;
Best Local Similarity 36.4%; Pred No. 4,2e-105;
Matches 302; Conservative 145; Mismatches 289; Indels 93; Gaps 22;

QY 8 VIDSSEYVNGKEPQVITIMGAENGERVILIDSRFPYFALAGADPKOV---AQRIR 63
DB 2 IIDDYITDEGKVIIRIFK-KENGEKPIEYDYEPIFYALKODSAIEVKKIIENRG 60
QY 64 ALSRPKSPITIGVEDDKKKTGCRPRVLRIRIVLPEAVREYRELVKNVGEVLEADIRP 123
DB 61 TVVTVKR---VEKVOKKFLGRPEVWKLYFTHPQDVPAIRDKIREHGAVIDIYEYDIP 116
QY 124 AMRYLIDHDLFPFTWYRVEAEPLNKGPRVDKYLVSRRPEPLYGEALAPTKLPDLRL 183
DB 117 AKRYLIDKGLV-----PWECD-----BELKML 138
QY 184 AFDEEVYSKQSPRPERDPVIVIAVKTDDDEVL-----FIAEGKDRKPIREPEY 235
DB 139 AFDIQTLVHEGEFPAE-GPILMISADEEGARVITWKVVDLPYDVVSTEREMIKRFLAV 197
QY 236 VRRYDPIIVGNNHNDMPYILRRARILIGIKLDVTRVGAEPSTVHG--VSVPGR 292
DB 198 VKEDPDLITVNGDNDPFAVLKRCCEKGIINFALGRD-GSEPKIQMKDRPRAVEYKRI 256
QY 293 NVDLYDAEEMPEIKISLEEVNAEYGVNKKSEYIIMWEIPDYDDPKRPLIQYAR 352
DB 257 HFDLIPVIRRTINLPYTLLEAVTEAVFGQPK-EKVYAE-EITPAWETGENLERVARYSM 313
QY 353 DDVATYGLAEKILIPALISYVTGLPLDQVAMSVEFRLEWYLIRAFKMKELVNNRYE 412
DB 314 EDKATYELGKELPMEQSLRILIGSLWDVSRSTGNLVEWFLRAQYERNELAPNKPD 373
QY 413 RPE----ETVGAIVLEPLRGVHENIAVLDPSSMYPNIMIKYVGPDLVRPGEKCEC 467
DB 374 EKELARRROSYEGGYVEPERGLMENIVYLDPRSLYPSIITHNVSPTLMRE----- 426
QY 468 GCWE---APEVGRFRRCPPGPFKTVLEHLELRKRVAEWKYPPDSPEYRLDEROKA 524
DB 427 GCEYDVAPQVGRFCDFPGFIPSLIGDLLEBOKIKKKWKA-TIDPIERKLDYRQRA 485
QY 525 LKVLNANASYGMGSGARWYCRECAKAVTAMGRHLIRPAL-NIARKLGLKVIYGTDSLF 583
DB 486 IKILANSYGYGYARARWYCKEACASVTAMGREYITMTIKEIBEKGFVYISDTDGF 545
QY 584 VTY---DPEKVE---NEIKIKEEL--GEFKLEKYKSLFFTEAKRYAGLLEBGRID 634
DB 546 ATIPGADAEIVKKGAMFELNVIYNAKLPGLALELEGEFYKRGFV-TKKYAVIDEBKIT 604
QY 635 IVGFPAVAGWCCELAKEVQTKVEIVLKTSEVNKAVEYVRKIVEKEGKVPLEKLVIRK 694
DB 605 TRGLEIVARDMSEIATKETAQARVLEALLKDGDEKAVRIKEVEKLSKEVEPPEKLVIE 664
QY 695 TLSKRLSEYTTAEHPVVAANKRMLSAGYRVSPGDKIGYIVYVKGSGRIISORAMPYFWKDP 754
DB 665 QITRDLKDYKATGPHVAVARLRAAGVKIRPGTVISYIVLKGSGRIDRAIP-FDEPDP 723
QY 755 --QIDVTVYVDHIIIPALALILGFGSITEKLLKASATGOKTLFPDLAKK 801
DB 724 KHKYDAEYIENQVLPAYERILRAPGYRKEDLRQKTRQVGLSAMLKPK 772

Search completed: April 23, 2003, 07:31:47
Job time : 33 secs